

UNIVERSIDADE FEDERAL DE SÃO PAULO
INSTITUTO DE CIÊNCIA E TECNOLOGIA
PÓS-GRADUAÇÃO EM BIOTECNOLOGIA

JULIANA APARECIDA DE MORAIS

**ANÁLISE DO REPERTÓRIO DE PROTEÍNAS SECRETADAS POR VIAS NÃO
CONVENCIONAIS EM SECRETOMAS TUMORAIS**

São José dos Campos
2021

UNIVERSIDADE FEDERAL DE SÃO PAULO
INSTITUTO DE CIÊNCIA E TECNOLOGIA
PÓS-GRADUAÇÃO EM BIOTECNOLOGIA

JULIANA APARECIDA DE MORAIS

**ANÁLISE DO REPERTÓRIO DE PROTEÍNAS SECRETADAS POR VIAS NÃO
CONVENCIONAIS EM SECRETOMAS TUMORAIS**

Dissertação apresentada ao Programa de Pós-Graduação em Biotecnologia, da Universidade Federal de São Paulo, como parte dos requisitos para obtenção do título de Mestre em Biotecnologia.

Orientador: Prof. Dr. André Zelanis Palitot Pereira.

São José dos Campos
2021

De Moraes, Juliana Aparecida

Análise do repertório de proteínas secretadas por vias não convencionais em secretomas tumorais / Juliana Aparecida De Moraes

Orientador(a) André Zelanis Palitot Pereira. - São José dos Campos, 2021.

112 p.

Dissertação (Mestrado - Programa de Pós-Graduação em Mestrado em Biotecnologia) - Universidade Federal de São Paulo - Instituto de Ciência e Tecnologia, 2021.

1. Secretoma. 2. Vias de secreção proteica não convencionais. 3. Câncer. 4. Bioinformática. 5. SUMOilação. I. Zelanis Palitot Pereira, André, orientador(a). II. Título.

UNIVERSIDADE FEDERAL DE SÃO PAULO
INSTITUTO DE CIÊNCIA E TECNOLOGIA
PÓS-GRADUAÇÃO EM BIOTECNOLOGIA

Chefe do Departamento: Prof. Dr. João Marcos Batista Júnior

Coordenador do curso de pós-graduação: Profa. Dra. Cláudia Barbosa Ladeira de
Campos

JULIANA APARECIDA DE MORAIS

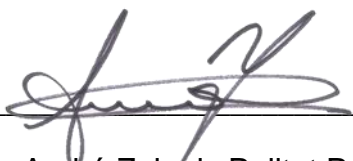
**ANÁLISE DO REPERTÓRIO DE PROTEÍNAS SECRETADAS POR VIAS NÃO
CONVENCIONAIS EM SECRETOMAS TUMORAIS**

Dissertação apresentada à Universidade Federal
São Paulo como requisito parcial para obtenção do
título de Mestre (a) em Ciências.

Área de Concentração: Biotecnologia

Aprovada em 22 de fevereiro de 2021.

Presidente da Banca:



Prof. Dr. André Zelanis Palitot Pereira

Banca Examinadora:

Profa. Dra. Cláudia Barbosa Ladeira de Campos

Prof. Dr. Giuseppe Palmisano

Prof. Dr. Milton Yutaka Nishiyama Junior

Dedicatória

*Dedico este trabalho a todos os pacientes com
câncer; vocês foram a minha única e verdadeira
motivação para o desenvolvimento deste trabalho.*

Agradecimentos

Agradeço primeiramente à Deus, sempre solícito e amoroso em Suas orientações, e por Seu acolhimento, principalmente nos momentos mais desafiadores. Agradeço também aos bons espíritos e ao meu mentor pelo amparo, aconselhamento e, sobretudo, por me ajudar a enxergar muito além do que eu mesma podia ver; reconheço que sem eles este trabalho jamais seria possível.

Agradeço aos meus pais pelo amor incondicional e pela liberdade necessária para que eu sempre pudesse buscar minha evolução neste plano terreno. À minha irmã e à minha sobrinha por sempre estarem ao meu lado.

Ao meu orientador, André Zelanis, pela paciência, confiança e por me orientar com primazia não apenas durante o mestrado, mas também durante minha graduação.

Ao Programa de Pós-Graduação em Biotecnologia da Universidade Federal de São Paulo e a todos os seus membros.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pelo apoio financeiro.

Aos colegas de laboratório, e em especial ao Tarcísio, pois sem os desdobramentos de sua pesquisa, a ideia deste trabalho provavelmente jamais teria sido cogitada.

Agradeço também minha querida amiga Sofia pela paciência e por todas as correções pontuais, além das inúmeras discussões sobre a biologia do câncer, sempre me fazendo olhar com novos olhos o desenrolar deste trabalho.

Por fim, agradeço a todos que direta ou indiretamente me auxiliaram no decorrer desta etapa e que, de algum modo, me ajudaram tanto na minha evolução profissional quanto pessoal, tornando essa caminhada mais leve e agradável.

A todos vocês, meu mais sincero muito obrigada!

*The boldness of asking deep questions may require unforeseen
flexibility if we are to accept the answers.*

Brian Greene

Resumo

Processos patológicos como o câncer têm um importante reflexo na síntese proteica e, consequentemente, no repertório de proteínas secretadas pelas células tumorais. No contexto da biologia tumoral, estas proteínas secretadas possuem um papel fundamental na comunicação intercelular. Há, inclusive, evidências de proteínas intracelulares fora da célula desempenhando diferentes funções no desenvolvimento e progressão do câncer, ressaltando a importância do estudo de vias de secreção proteica não convencionais. Deste modo, linhas de pesquisa em oncologia têm explorado cada vez mais o estudo de proteínas secretadas para o entendimento da interação de células tumorais com o microambiente. **Objetivos:** identificação do repertório de proteínas secretadas não convencionalmente nos secretomas de câncer de mama, melanoma, cólon, ovário e sarcoma de *Ewing*. **Métodos:** arquivos de espectrômetros de massas oriundos de diferentes tipos de secretomas tumorais – cólon, mama, melanoma, ovário e sarcoma de *Ewing* – foram obtidos da plataforma *ProteomeXchange*. As buscas no banco de dados foram feitas usando a plataforma *Trans Proteomics Pipeline* e o algoritmo de busca *Comet*, considerando uma taxa de falsos positivos (FDR) $\leq 1\%$ ao nível de proteína/peptídeo. Filtros foram aplicados nas listas de proteínas através de anotações manuais no banco de dados UniProtKB/Swiss-Prot. Análises de enriquecimento funcional/localização celular e enriquecimento de vias foram feitas utilizando as plataformas *Cytoscape* e *Reactome*. Análises mutacionais foram realizadas utilizando dados de pacientes, disponíveis na plataforma *cBioPortal for Cancer Genomics* e a avaliação da expressão de proteínas selecionadas em amostras de tecido tumorais e não tumorais foi realizada na plataforma *The Human Protein Atlas*. A análise dos dados e a criação de imagens foram feitas através de *scripts* próprios criados a partir da linguagem de programação R. **Resultados:** 6.092 proteínas foram identificadas. Destas, 2.312 proteínas não pertenciam a nenhuma via de secreção convencional e foram selecionadas para análise. A porcentagem de proteínas possivelmente de via secretória não canônica nos secretomas em relação aos secretomas pré-filtragem variou de 17% (melanoma) a 39% (ovário) e mostrou uma similaridade proteica dentre diferentes tipos de secretomas tumorais como, por exemplo, cólon e ovário. Análises de enriquecimento das dezenove proteínas compartilhadas em todos os secretomas estudados mostraram que elas estão enriquecidas no núcleo e sua principal função está relacionada à SUMOilação, uma modificação pós-traducional, de proteínas. A análise de amostras de tecido de pacientes na plataforma *The Human Protein Atlas* corrobora com a mudança de localização do núcleo para o citoplasma de três proteínas, dentre as dezenove compartilhadas, no câncer de cólon, ovário e melanoma, indicando a presença de proteínas nucleares no secretoma. **Conclusão:** foi possível observar o enriquecimento de proteínas intracelulares, sobretudo de proteínas nucleares – inclusive relacionadas ao transporte nuclear – no repertório de proteínas secretadas por diferentes tipos de câncer, característica que sugere vias não-canônicas sendo acionadas por estas proteínas. Portanto, a principal hipótese é de que processos de SUMOilação podem estar envolvidos na exportação e secreção das proteínas nucleares.

Palavras-chave: Secretoma. Vias de secreção proteica não convencionais. Câncer. Bioinformática. SUMOilação.

Abstract

Pathological processes such as cancer have an important effect on protein synthesis and, consequently, on the repertoire of proteins secreted by tumor cells. In the context of tumor biology, these secreted proteins play an essential role in intercellular communication. There is even evidence of intracellular proteins outside the cell performing different functions in the development and progression of cancer, highlighting the importance of studying non-conventional protein secretion pathways. Thus, research in oncology have increasingly explored the study of secreted proteins to understand the interaction of tumor cells with the microenvironment. **Aim:** identify the repertoire of proteins unconventionally secreted in the secretomes of breast cancer, melanoma, colon, ovary, and Ewing's sarcoma. **Methods:** files of mass spectrometers from different types of tumor secretomes – colon, breast, melanoma, ovary, and Ewing's sarcoma – were obtained from the ProteomeXchange platform. The database searches were performed using the Trans Proteomics Pipeline platform and the Comet search algorithm, considering a false positive rate (FDR) $\leq 1\%$ at the protein/peptide level. Filters were applied to the protein lists through manual annotations in the UniProtKB/Swiss-Prot database. We performed functional enrichment/cell location and pathway enrichment analyzes using the Cytoscape and Reactome platforms. After, we performed mutational analyzes using patient data, available on the cBioPortal for Cancer Genomics platform and the evaluation of the expression of selected proteins in tumor and non-tumor tissue samples on The Human Protein Atlas platform. We performed data analysis and created images using in-house scripts created with R programming language. **Results:** We identified 6,092 proteins. Of these, 2,312 proteins did not belong to any conventional secretion pathway and were selected for further analysis. The percentage of proteins possibly from non-canonical secretory pathway in the secretomes (when compared to those in the pre-filtering step) ranged from 17% (melanoma) to 39% (ovary) and showed a similarity among different types of tumor secretomes, such as colon and ovary. Enrichment analysis of the nineteen shared proteins in all secretomes showed that they are enriched in the nucleus and their main function is related to SUMOylation, a post-translational modification. The analysis of tissue samples from patients on The Human Protein Atlas platform corroborates with the change of nucleus location to the cytoplasm of three proteins, among the nineteen shared ones, in colon cancer, ovary cancer, and which indicates the presence of nuclear proteins in the secretomes. **Conclusion:** we observed the enrichment of intracellular proteins, especially nuclear proteins – including those related to nuclear transport – in the repertoire of proteins secreted by different types of cancer, which suggests non-canonical pathways being triggered by these proteins. Therefore, the main hypothesis is that SUMOylation processes may be involved in the export and secretion of nuclear proteins.

Keywords: Secretome. Unconventional protein secretion pathways. Cancer. Bioinformatics. SUMOylation.

Lista de figuras

Figura 1 – Roteiro simplificado do tráfego de proteínas em uma célula eucariótica....	4
Figura 2 – Sequências-sinal típicas de acordo com sua função.....	5
Figura 3 – Estrutura do envelope nuclear	6
Figura 4 – Ran-GDP e Ran-GTP no transporte nuclear	8
Figura 5 – Direcionalidade na importação e exportação nuclear.....	9
Figura 6 – Vias secretoras constitutiva e regulada.....	11
Figura 7 – Visão geral da via de secreção proteica convencional.....	13
Figura 8 – Visão geral das vias de secreção proteica convencional e não convencionais.....	15
Figura 9 – CELA3 em amostra histológica no pâncreas	16
Figura 10 – A via SUMO	20
Figura 11 – Visão geral do processamento e análise dos dados avaliados nesse projeto	23
Figura 12 – Critérios estabelecidos em cada etapa de filtragem para posterior identificação e remoção das proteínas.....	26
Figura 13 – Perfil geral do repertório das proteínas totais e de cada tipo de câncer.....	32
Figura 14 – Composição dos secretomas pós-filtragem	34
Figura 15 – Visão geral das vias enriquecidas para as proteínas únicas.....	36
Figura 16 – Enriquecimento do componente celular por secretoma tumoral	37
Figura 17 – Visão geral das proteínas compartilhadas com maior número de intersecções entre os diferentes tipos de câncer	39
Figura 18 – Análise de enriquecimento do componente celular das proteínas compartilhadas	40
Figura 19 – Panorama geral das proteínas compartilhadas dentre os diferentes secretomas tumorais	42
Figura 20 – Proteínas exclusivas por tipo de câncer.....	43
Figura 21 – Perfil das alterações genéticas dos dezenove genes compartilhados em pacientes de acordo com os diferentes tipos de câncer estudados	44
Figura 22 – Perfil mutacional dos genes LMNA, NPM1, NUDC e STIP1 em pacientes com câncer.....	45
Figura 23 – Localização das proteínas compartilhadas em tecidos não tumorais e tumorais.....	49

Lista de tabelas

Tabela 1 – Predição do secretoma humano por quatro métodos de predição diferentes.....	17
Tabela 2 – Arquivos selecionados no <i>ProteomeXChange Consortium</i>	25
Tabela 3 – Palavras-chave buscadas no banco de dados em cada etapa dos filtros para identificação e remoção das proteínas.....	27
Tabela 4 – Dados tumorais de pacientes obtidos na plataforma <i>cBioPortal</i>	30
Tabela 5 – Principais vias enriquecidas por tipo de câncer.....	38
Tabela 6 – Principais vias enriquecidas nas dezenove proteínas compartilhadas	41
Tabela 7 – Proteínas nucleares encontradas compartilhadas nos secretomas estudados e suas respectivas características em organismo saudável.	47

Lista de abreviaturas e siglas

cBioPortal	<i>cBioPortal for Cancer Genomics</i>
CELA3A	<i>Chymotrypsin-like elastase family, member 3A</i>
CK2	<i>Casein kinase 2</i>
COPII	<i>Coat protein II</i>
FDR	<i>False Discovery Rate</i>
GAP	<i>GTPase-activating protein</i>
GEF	<i>Guanine exchange factor</i>
GO	<i>Gene Ontology</i>
HMGA1	<i>High mobility group A1</i>
HSP	<i>Heat shock protein</i>
ID3	<i>Inhibitor of DNA binding 3</i>
LC-MS/MS	<i>Cromatografia líquida acoplada à espectrometria de massas em tandem</i>
LMA	<i>Leucemia mieloide aguda</i>
MPT	<i>Modificação pós-traducional</i>
NLS	<i>Nuclear localization signal</i>
NoLS	<i>Nucleolar localization signal</i>
NPC	<i>Nuclear pore complex</i>
NPM1	<i>Nucleophosmin 1</i>
NPM1c	<i>NPM1 citoplasmática</i>
PrP ^C	<i>Proteína príon celular</i>
RanGAP1	<i>Ran GTPase-activating protein 1</i>
RanBP1	<i>Ran binding protein 1</i>
RanBP2	<i>Ran binding protein 2</i>
RE	<i>Retículo endoplasmático rugoso</i>
SAE1	<i>SUMO-activating enzyme subunit 1</i>

STIP1	<i>Stress-induced phosphoprotein 1</i>
SUMO	<i>Small ubiquitin-like modifier</i>
TCGA	<i>The Cancer Genome Atlas</i>
TGN	<i>Trans Golgi network</i>
TPP	<i>Trans-Proteomic Pipeline</i>
UBA2	<i>Ubiquitin-like 1-activating enzyme E1B</i>
UPS	<i>Unconventional protein secretion</i>

SUMÁRIO

Dedicatória.....	vi
Agradecimentos	vii
Resumo	ix
<i>Abstract</i>	x
Lista de figuras.....	xi
Lista de tabelas	xii
Lista de abreviaturas e siglas	xiii
1 INTRODUÇÃO	1
2 REVISÃO DA BIBLIOGRAFIA	3
2.1 Endereçamento e tráfego de proteínas	3
2.2 Transporte de proteínas nucleares.....	5
2.2.1 Sinais de endereçamento.....	5
2.2.2 Direcionalidade no transporte nuclear	7
2.3 Via de secreção proteica convencional (ou clássica).....	10
2.3.1 Do retículo endoplasmático até Golgi.....	10
2.3.2 Transporte da rede <i>trans</i> de Golgi para o exterior da célula	11
2.4 Vias de secreção proteica não convencionais	13
2.5 O secretoma.....	15
2.6 Secretoma e câncer	17
2.7 SUMOilação	18
3 Objetivos	22
4 MATERIAL E MÉTODOS	23
4.1 Aquisição dos dados de secretoma	24
4.2 Processamento dos arquivos brutos	25
4.3 Filtragem das proteínas.....	26
4.4 Processamento e análise das proteínas possivelmente de via secretória não canônica	29
4.4.1 Análise de intersecções proteicas e estudo de mutações	29
4.4.2 Validação da localização das proteínas compartilhadas nucleares através de amostras de tecido.....	30
5 RESULTADOS.....	31
5.1 Repertório das proteínas do secretoma pós-filtragem	31
5.2 Análise das proteínas compartilhadas nos diferentes secretomas tumorais	39
5.2.1 Proteínas compartilhadas e análises de enriquecimento	39
5.2.2 Alterações genéticas e perfil mutacional	43
5.3 Localização biológica das proteínas nucleares compartilhadas em amostras histológicas.....	46
6 DISCUSSÃO	53
6.1 Análise do repertório geral e excedente proteicos nos secretomas estudados	53
6.2 O papel da SUMOilação em diferentes proteínas	55

6.3 Evidências de proteínas intracelulares secretadas e seu novo papel em patogêneses	57
6.4 Considerações finais	63
7 CONCLUSÕES	66
8 DISPONIBILIDADE DOS DADOS	68
REFERÊNCIAS.....	69
ANEXOS	82
A – Visão geral dos arquivos selecionados e ferramentas utilizadas pelos respectivos autores nas buscas	83
B – Características específicas das amostras de câncer de cólon selecionadas.....	84
C – Características específicas das amostras de câncer de mama selecionadas.....	86
D – Características específicas das amostras de melanoma selecionadas.....	87
E – Características específicas das amostras de câncer de ovário selecionadas.....	88
F – Características específicas das amostras de sarcoma de Ewing selecionadas.....	90
G – Identificação no Uniprot/Swissprot das proteínas mitocondriais e relacionadas à apoptose removidas.....	91
H – Lista dos genes das cem principais proteínas que são frequentemente identificadas no Vesiclepedia e ExoCarta.....	92
I – Vias enqueridas mais relevantes referentes às proteínas únicas.....	93
J – Sítios de SUMOilação identificados nas proteínas compartilhadas	94
L – Descrição da localização das proteínas compartilhadas em tecidos no <i>The Protein Atlas</i>	96

1 INTRODUÇÃO

A síntese de todas as proteínas ocorre através dos ribossomos, presentes no citosol de células eucarióticas; a exceção ocorre devido a proteínas sintetizadas em mitocôndrias e plastídios (ALBERTS, B.; BRAY, D.; LEWIS, J.; RAFF, M.; ROBERTS, K.; WATSON, 2017). O destino de cada proteína sintetizada depende de sua sequência de aminoácidos, que funciona como um código de endereçamento, indicando o local para onde a proteína recém-sintetizada deverá ir, seja dentro ou fora da célula; contudo, algumas proteínas não possuem qualquer sinal de endereçamento e permanecem no citosol como proteínas residentes (VON HEIJNE, 1990).

Um exemplo são os sinais de endereçamento de proteínas nucleares, chamados de sinais de localização nuclear, que podem estar situados em qualquer local da estrutura primária de proteínas (QUIMBY; CORBETT, 2001). Para iniciar a importação nuclear, estes sinais de endereçamento devem ser reconhecidos por receptores de importação nuclear, também chamadas de importinas; esta ligação pode ser direta ou indireta, através de transportadores. Estes receptores se ligam tanto aos sinais de endereçamento das proteínas nucleares quanto às sequências de fenil alanina-glicina presentes na região central do poro nuclear e assim são guiados até chegarem ao núcleo, dissociando-se da proteína nuclear e retornando então ao citosol (QUIMBY; CORBETT, 2001).

As proteínas destinadas à secreção fazem parte de uma via de secreção proteica convencional ou clássica. Elas são encaminhadas primeiramente do citosol ao retículo endoplasmático; estas proteínas só deixam o retículo endoplasmático após estarem completamente enoveladas e então seguem empacotadas em vesículas, e posteriormente agrupamentos, para o complexo de Golgi, onde sofrem uma série de modificações conforme avançam pelas redes *cis* e *trans* de Golgi, quando são, por fim, distribuídas para vesículas secretoras ou lisossomos (HERRMANN; MALKUS; SCHEKMAN, 1999). Após carregadas, estas vesículas chegam ao local de secreção e esperam, próximas a membrana plasmática, um sinal para que secretem seu conteúdo (VIOTTI, 2016).

Embora a via de secreção proteica convencional seja responsável pela secreção da maioria das proteínas (VIOTTI, 2016), não é incomum o uso de vias alternativas de secreção proteica que não obedeçam a ordem retículo endoplasmático

-Golgi. Estas vias são chamadas de *vias de secreção proteica não convencionais* e são induzidas por diferentes tipos de estresse às células, tais como o estresse oriundo de doenças inflamatórias e câncer (NICKEL; RABOUILLE, 2009; RABOUILLE, 2017).

Deste modo, processos patológicos como o câncer têm um importante reflexo na síntese proteica e, conseqüentemente, no repertório de proteínas secretadas pelas células tumorais (XUE; LU; LAI, 2008). No contexto da biologia tumoral, estas proteínas secretadas possuem um papel fundamental na comunicação intercelular, havendo inclusive relatos da presença de proteínas intracelulares fora da célula desempenhando diferentes funções no desenvolvimento e progressão tumoral (MA *et al.*, 2020; MÉNDEZ *et al.*, 2019; ODUNUGA; LONGSHAW; BLATCH, 2004). Deste modo, linhas de pesquisa em oncologia têm explorado cada vez mais o estudo de proteínas secretadas para o entendimento da interação de células tumorais com o microambiente (BROWN *et al.*, 2012; LIN *et al.*, 2013; SANTOS; MARTINS; HAJJ, 2017; VILLARREAL *et al.*, 2013).

2 REVISÃO DA BIBLIOGRAFIA

2.1 Endereçamento e tráfego de proteínas

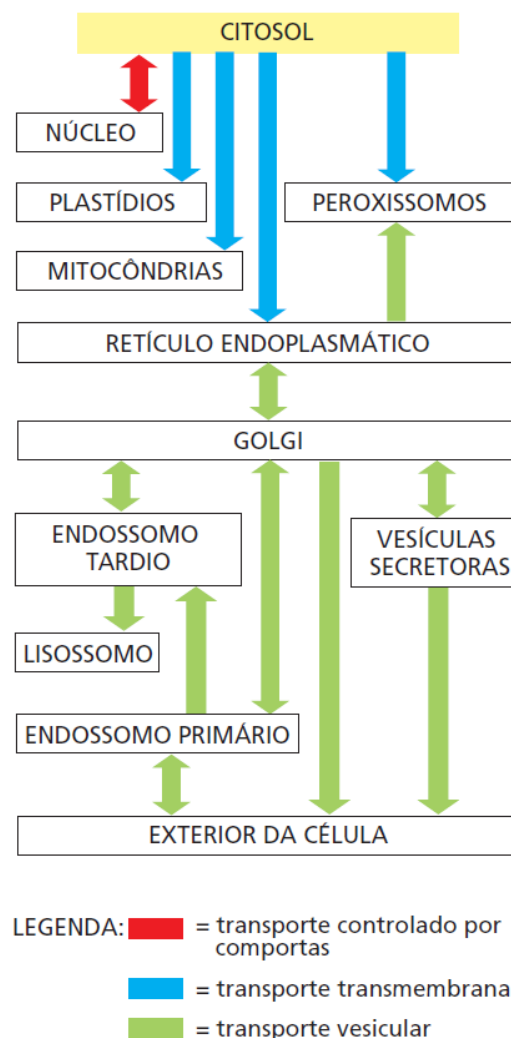
A síntese de todas as proteínas começa no citosol, através de ribossomos; a exceção ocorre em poucas proteínas que são sintetizadas nos ribossomos de mitocôndrias e plastídios. O destino posterior de cada proteína depende de sua sequência de aminoácidos, que podem abranger sinais de endereçamento que direcionam o envio da proteína para locais fora do citosol ou então para a superfície de outras organelas. Algumas proteínas não possuem qualquer sinal de endereçamento e, portanto, permanecem no citosol como proteínas residentes; outras apresentam sinais de endereçamento que são específicos para o seu direcionamento do citosol ao núcleo, ao retículo endoplasmático rugoso (RE), aos plastídios ou aos peroxissomos. Além disso, os sinais de endereçamento também podem conduzir o transporte de proteínas do RE a outros destinos dentro da célula (OTT; AMUNTS; BROWN, 2016; PALADE, 1975)

Os princípios gerais aos quais os sinais de endereçamento atuam em uma célula eucariótica são compreendidos por três mecanismos fundamentais: transporte controlado por comportas; transporte transmembrana ou translocação de proteínas; e transporte vesicular (Figura 1). No transporte controlado por comportas, as proteínas e moléculas de RNA se movem entre citosol e núcleo através de complexos do poro nuclear no envelope nuclear; estes complexos funcionam como canais seletivos e auxiliam tanto o transporte ativo de macromoléculas específicas, quanto a difusão livre de pequenas moléculas. Na translocação de proteínas, translocadores de proteínas transmembrana transportam diretamente proteínas específicas através da membrana do citosol para outro espaço com topologia diferente; para passar pelo translocador, a proteína transportada geralmente precisa se desdobrar. No transporte vesicular, vesículas de transporte esféricas de diferentes tamanhos levam proteínas de um espaço para outro. A transferência de proteínas solúveis do RE ao complexo de Golgi ocorre dessa maneira (ALBERTS, B.; BRAY, D.; LEWIS, J.; RAFF, M.; ROBERTS, K.; WATSON, 2017; PALADE, 1975; QUIMBY; CORBETT, 2001).

Cada uma dessas formas de transferência de proteínas é guiada por sinais de endereçamento na proteína transportada, que são reconhecidos pelos receptores de endereçamento complementares. Se uma proteína deve ser transportada para o

núcleo, por exemplo, seu sinal deve ser reconhecido por receptores ao longo da sua trajetória até o poro nuclear; o mesmo ocorre para os outros tipos de transporte; se uma proteína for carregada por uma vesícula, um receptor na membrana apropriada deve reconhecer seu sinal de endereçamento. Estes receptores de endereçamento reconhecem **classes** de proteínas mais do que proteínas específicas, e podem ser vistos como sistemas de transporte que são dedicados ao transporte de inúmeros componentes distintos até seu destino correto dentro da célula (ALBERTS, B.; BRAY, D.; LEWIS, J.; RAFF, M.; ROBERTS, K.; WATSON, 2017).

Figura 1 – Roteiro simplificado do tráfego de proteínas em uma célula eucariótica



Fonte: Extraído e adaptado ALBERTS *et al.*, 2017, p. 646

Nota. Roteiro simplificado do tráfego de proteínas em uma célula eucariótica. As proteínas podem se mover de um setor a outro da célula através de comportas (vermelho), translocação (azul) ou transporte vesicular (verde). Os sinais de endereçamento direcionam o caminho da proteína ao longo dos compartimentos (retângulos) e para muitas proteínas, o direcionamento só termina quando as proteínas alcançam seu destino.

Em geral, os sinais de endereçamento das proteínas envolvidos no transporte transmembrana encontram-se em geral, em um trecho de aminoácidos de 15 a 60 resíduos. Estas sequências-sinal são frequentemente encontradas na porção N-terminal da cadeia polipeptídica e, em diversos casos, peptidases-sinal específicas removem a sequência-sinal da proteína, indicando que o transporte está finalizado. Em outros casos, como nas proteínas nucleares, o sinal de endereçamento encontra-se em extensões internas e permanecem como parte da proteína, não sendo removido pelas peptidases-sinal (PALADE, 1975; QUIMBY; CORBETT, 2001).

Cada sequência-sinal define um destino específico na célula. As proteínas destinadas ao RE geralmente possuem uma sequência-sinal na sua região N-terminal, e uma de suas características é uma sequência composta de 5 a 10 aminoácidos hidrofóbicos. Muitas dessas proteínas se deslocarão do RE rugoso para Golgi, mas há algumas com um sinal específico de quatro aminoácidos em sua região C-terminal que identifica as proteínas como residentes no RE e assim, elas retornam a ele (VON HEIJNE, 1990). A Figura 2 apresenta algumas dessas sequências-sinal específicas.

Figura 2 – Sequências-sinal típicas de acordo com sua função

Função da sequência-sinal	Exemplo de sequência-sinal
Importar para o núcleo	-Pro-Pro- Lys-Lys-Lys-Arg-Lys -Val-
Exportar do núcleo	- Met -Glu-Glu- Leu -Ser-Gln-Ala- Leu -Ala-Ser-Ser- Phe -
Importar para a mitocôndria	*H ₃ N-Met-Leu-Ser-Leu- Arg -Gln-Ser-Ile- Arg -Phe-Phe- Lys -Pro-Ala-Thr- Arg -Thr-Leu-Cys-Ser-Ser- Arg -Tyr-Leu-Leu-
Importar para o plastídio	*H ₃ N-Met-Val-Ala-Met-Ala-Met-Ala- Ser -Leu-Gln- Ser-Ser -Met- Ser-Ser -Leu- Ser -Leu- Ser-Ser -Asn-Ser-Phe-Leu-Gly-Gln-Pro-Leu- Ser -Pro-Ile- Thr -Leu- Ser -Pro-Phe-Leu-Gln-Gly-
Importar para os peroxissomos	- Ser-Lys-Leu -COO ⁻
Importar para o RE	*H ₃ N-Met-Met-Ser-Phe-Val-Ser- Leu-Leu-Leu-Val -Gly-Ile-Leu- Phe-Trp -Ala-Thr-Glu-Ala-Glu-Gln-Leu-Thr- Lys -Cys- Glu -Val-Phe-Gln-
Retornar ao RE	- Lys-Asp-Glu-Leu -COO ⁻

Alguns aspectos característicos das diferentes classes de sequências-sinal estão destacados em cores diferentes. Quando sua importância para a função da sequência-sinal é conhecida, os aminoácidos carregados positivamente são mostrados em *vermelho*, e os carregados negativamente, em *verde*. Do mesmo modo, os aminoácidos hidrofóbicos importantes são mostrados em *laranja*, e os aminoácidos hidroxilados, em *azul*. *H₃N indica a região N-terminal de uma proteína; COO⁻ indica a região C-terminal.

Fonte: Extraído e adaptado de ALBERTS *et al.*, 2017, p. 648

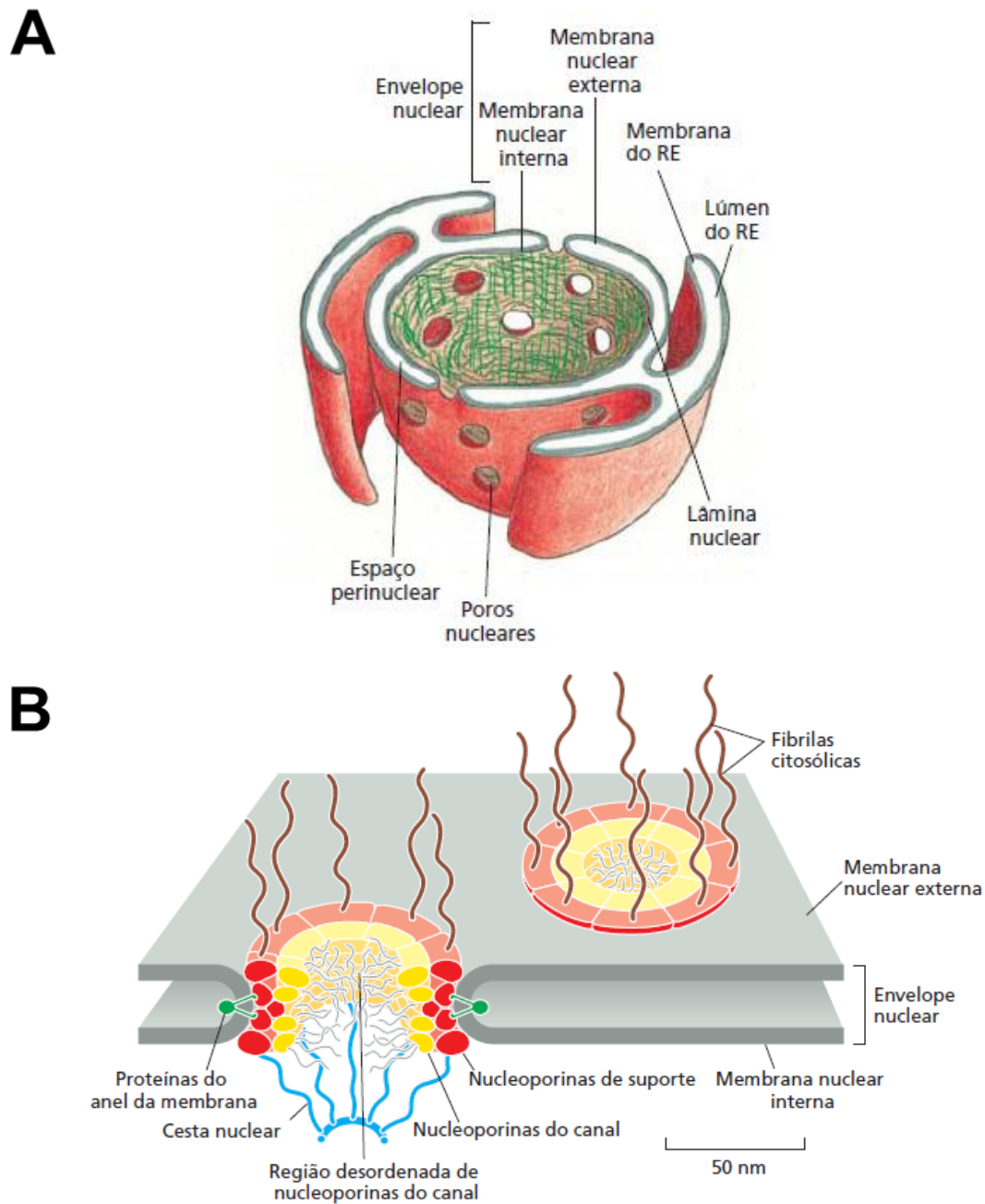
2.2 Transporte de proteínas nucleares

2.2.1 Sinais de endereçamento

Sinais de endereçamento das proteínas nucleares, chamados de sinais de localização nuclear (NLSs; do inglês, *nuclear localization signals*) podem estar situados em qualquer lugar na sequência de aminoácidos e, aparentemente, formam

alças na superfície da proteína. O transporte de proteínas nucleares ocorre através de complexos do poro nuclear (NPCs; do inglês, *nuclear pore complexes*), posicionados nos poros do núcleo, que estão presentes no envelope nuclear em todas as células eucarióticas (Figuras 3A; 3B) (HETZER; WENTE, 2009; ROTHBALLER; KUTAY, 2013)

Figura 3 – Estrutura do envelope nuclear



Fonte: Extraído e adaptado de ALBERTS *et al.*, 2017, p. 650-51

Nota. Envelope nuclear. (A) Estrutura do envelope nuclear, onde estão contidos os poros nucleares. (B) Estrutura dos complexos do poro nuclear (NPC) que se encontram posicionados no poro nuclear. Cada NPC é composto por um grupo de cerca de 30 proteínas diferentes, ou nucleoporinas; as proteínas que formam a porção central do NPC estão orientadas simetricamente ao longo do envelope nuclear de modo que tanto o lado nuclear quanto o lado citosólico pareçam idênticos.

O transporte pelos NPCs é diferente do transporte de proteínas transmembrana, pois ocorre através de um poro aquoso ao invés do uso de uma proteína transportadora e, por esta razão, as proteínas nucleares podem ser transportadas para dentro do núcleo, através de um NPC, em sua conformação completamente enovelada; todavia, as proteínas não nucleares precisam ser extensivamente desenoveladas durante seu transporte para outras organelas.

Para iniciar a importação nuclear, a maioria dos NLSs devem ser reconhecidos por receptores de importação nuclear, também chamados de importinas, através de uma ligação direta ou através de proteínas adaptadoras; alguns receptores de importação nuclear correlacionam-se estruturalmente com proteínas adaptadoras, indicando uma origem evolutiva em comum.

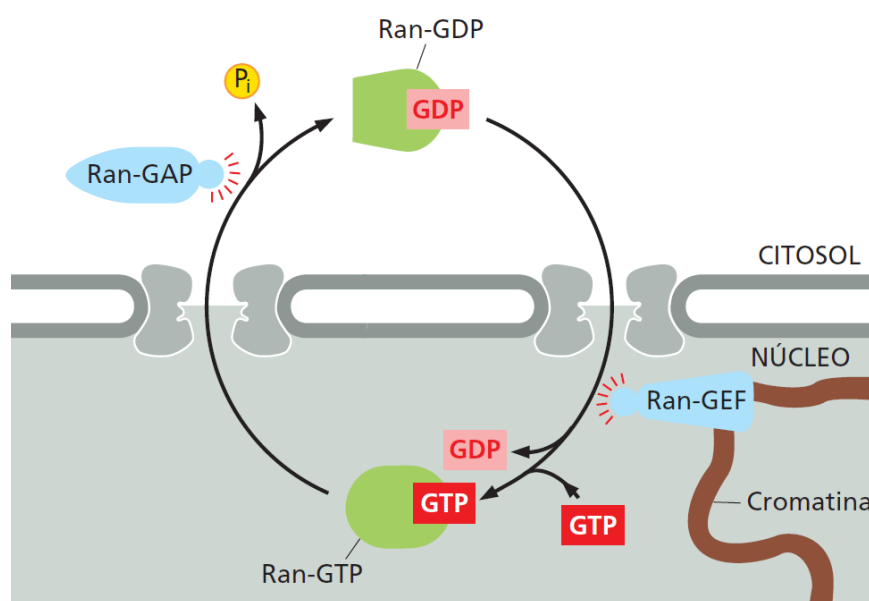
Estes receptores de importação se ligam tanto ao NLS da proteína a ser transportada quanto às sequências repetidas de fenilalanina–glicina (FG) nos domínios do canal de nucleoporinas, alinhados na região central do poro; essas repetições FG interagem fracamente e impõem uma barreira de permeabilidade a grandes moléculas, além de servirem como um local de ancoragem para os receptores nucleares de importação. Os receptores de importação se ligam às repetições FG ao longo do caminho e quando chegam no núcleo, dissociam-se da proteína e retornam ao citosol; sabe-se que a importação através dos NPCs concentra proteínas somente do núcleo, aumentando a ordem da célula (HÜLSMANN; LABOKHA; GÖRLICH, 2012; KÖHLER; HURT, 2007; STRAMBIO-DE-CASTILLIA; NIEPEL; ROUT, 2010; TRAN; WENTE, 2006)

2.2.2 Direcionalidade no transporte nuclear

Tanto a importação quanto a exportação nuclear dependem do armazenamento de energia em gradientes de concentração na forma ligada ao GTP da GTPase Ran monomérica. Assim como outras GTPases, Ran pode existir em dois estados conformacionais e a conversão entre estes dois estados é desencadeada por duas proteínas reguladoras Ran-específicas: uma proteína que ativa GTPase no citosol

(GAP; do inglês, *GTPase-activating protein*) e aciona a hidrólise de GDP, convertendo Ran-GTP em Ran-GDP, e um fator de troca de guanina (GEF; do inglês, *guanine exchange factor*) nuclear, que promove a troca de GDP para GTP, convertendo Ran-GDP em Ran-GTP. Nota-se que enquanto o citosol contém principalmente Ran-GDP, o núcleo contém sobretudo Ran-GTP (Figura 4).

Figura 4 – Ran-GDP e Ran-GTP no transporte nuclear



Fonte: Extraído e adaptado de ALBERTS *et al.*, 2017, p. 653.

Nota. Direção do transporte nuclear a partir de Ran-GDP e Ran-GTP. Nota-se que enquanto o citosol contém principalmente Ran-GDP, o núcleo contém sobretudo Ran-GTP.

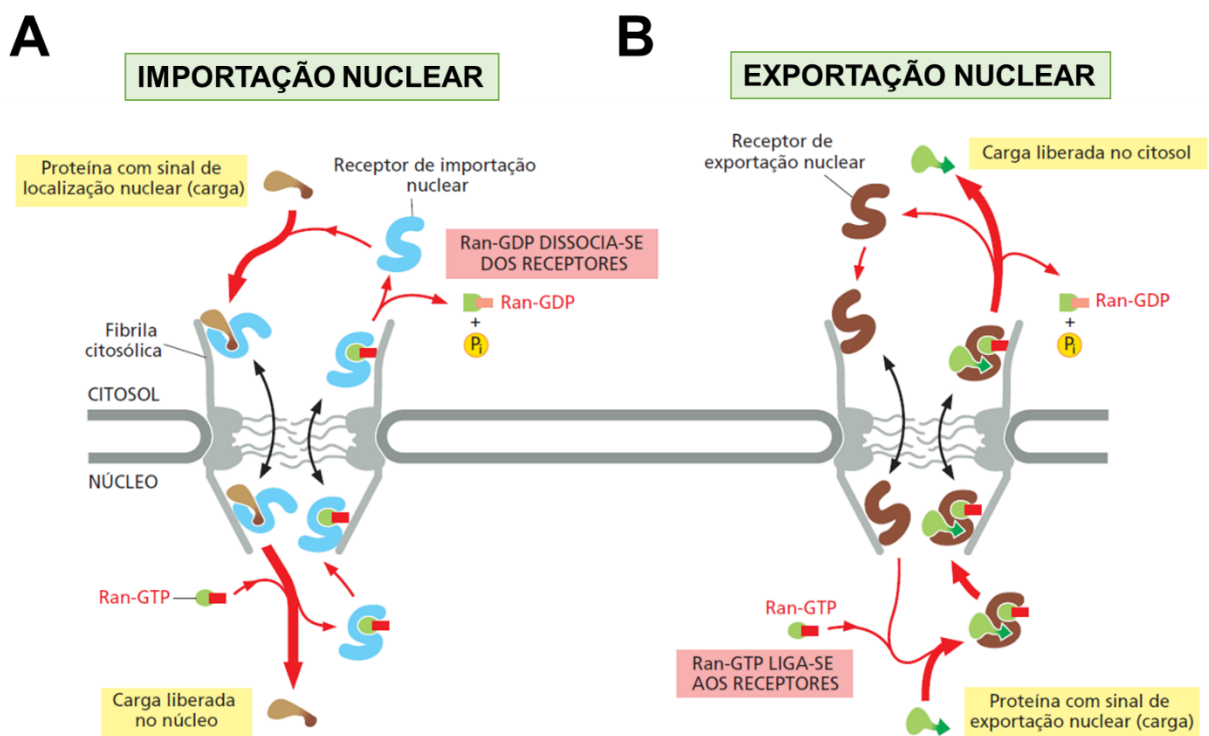
O gradiente das duas formas conformacionais de Ran dirige o transporte nuclear na direção correta. A junção dos receptores de importação nuclear nas repetições FG no lado citosólico do NPC ocorre somente quando estes receptores estão ligados à carga proteica apropriada, logo após eles entram no canal. Ao atingirem o lado nuclear do complexo do poro, Ran-GTP liga-se a eles e esta ligação faz os receptores de importação liberarem sua carga. No citosol, Ran-GDP não se liga a receptores de importação ou exportação e o descarregamento ocorre apenas no lado nuclear do NPC. Dessa maneira, a localização nuclear de Ran-GTP cria a direcionalidade do processo de importação.

Depois de descarregar sua carga no núcleo, o receptor de importação vazio com Ran-GTP é transportado de volta ao citosol através do complexo do poro, onde

Ran-GAP estimula Ran-GTP a hidrolisar seu GTP ligado, convertendo-o, assim, em Ran-GDP, o qual dissocia-se do receptor. Novamente no citosol, o receptor está pronto para outro ciclo de importação nuclear (Figura 5A).

Muito embora a exportação nuclear ocorra através de um mecanismo similar, Ran-GTP no núcleo promove a ligação da carga ao receptor de exportação, ao invés de promover a dissociação da mesma. Como o receptor de exportação se movimenta através do poro nuclear para o citosol, ele se liga a Ran-GAP que induz o receptor a hidrolisar seu GTP a GDP; logo após, o receptor de exportação libera tanto sua carga quanto RAN-GDP no citosol. Os receptores de exportação livres retornam ao núcleo para completar o ciclo (Figura 5B).

Figura 5 – Direcionalidade na importação e exportação nuclear



Fonte: Extraído e adaptado de ALBERTS *et al.*, 2017, p. 654.

Nota. Direção do transporte nuclear. Na figura, há a elucidação da direção da (A) importação e (B) exportação nuclear através de Ran-GDP e Ran-GTP.

2.3 Via de secreção proteica convencional (ou clássica)

2.3.1 Do retículo endoplasmático até Golgi

A maior parte da síntese de proteínas celulares ocorre na superfície citosólica do RE rugoso e as proteínas destinadas à secreção ou a outras organelas são importadas, primeiramente, do citosol para o RE. Já no lúmen do RE, essas proteínas se enovelam e se oligomerizam, de modo que só deixam o RE quando estão adequadamente enoveladas; as proteínas que não se enovelam ou oligomerizam corretamente são levadas de volta ao citosol, onde são degradadas em proteassomos (BRAAKMAN; BULLEID, 2011; BRODSKY; SKACH, 2011; VON HEIJNE, 2011; WALTER; RON, 2011).

Após serem corretamente enoveladas e montadas no RE, as proteínas são então empacotadas em vesículas de transporte revestidas por COPII (do inglês, *coat protein II*) que se destacam da membrana do RE; em seguida, estas vesículas perdem seu revestimento e se fundem umas às outras para formarem agrupamentos tubulares de vesículas. Em células animais, esses agrupamentos se movem para o complexo de Golgi, onde se fusionam uns aos outros para formar a rede *cis* de Golgi. Neste momento, proteínas que sejam residentes do RE e esquivaram-se de lá são devolvidas por esses agrupamentos por um transporte retrógrado.

O complexo de Golgi realiza uma série de modificações nas proteínas – por exemplo, processos de fosforilação e glicosilação – que recebe do RE e então as distribui para vesículas secretoras, membrana plasmática e lisossomos. O complexo de Golgi é uma organela que consiste em uma ou mais pilhas de cisternas em forma de disco e cada pilha é organizada com compartimentos funcionalmente distintos e estações especiais de seleção que são chamadas de rede *cis* de Golgi e rede *trans* de Golgi (TGN; do inglês, *trans Golgi network*). As proteínas se movem através das pilhas de Golgi em uma direção *cis–trans*. As novas proteínas concluídas terminam na rede *trans* de Golgi, que as empacota em vesículas de transporte para enfim despachá-las a seus destinos dentro ou fora da célula (EMR *et al.*, 2009; FARQUHAR; PALADE, 1998; LADINSKY *et al.*, 1999; PFEFFER, 2010; VARKI, 2011).

2.3.2 Transporte da rede *trans* de Golgi para o exterior da célula

As vesículas secretoras formam-se a partir da TGN e liberam seus conteúdos ao exterior da célula por exocitose em respostas a sinais específicos. O produto secretado pode ser tanto uma pequena molécula ou então uma proteína.

Há duas vias de secreção: vias de secreção constitutiva e regulada, sendo que as duas vias são oriundas da TGN. A via constitutiva secreta continuamente proteínas solúveis através de exocitose e fornece lipídeos e proteínas recém-sintetizados para a membrana plasmática. A via de secreção regulada ocorre através da seleção de proteínas na TGN, as quais são desviadas para as vesículas secretoras, que posteriormente esperam um sinal para liberarem seu conteúdo (Figura 6) (ALBERTS, B.; BRAY, D.; LEWIS, J.; RAFF, M.; ROBERTS, K.; WATSON, 2017; BURGESS; KELLY, 1987).

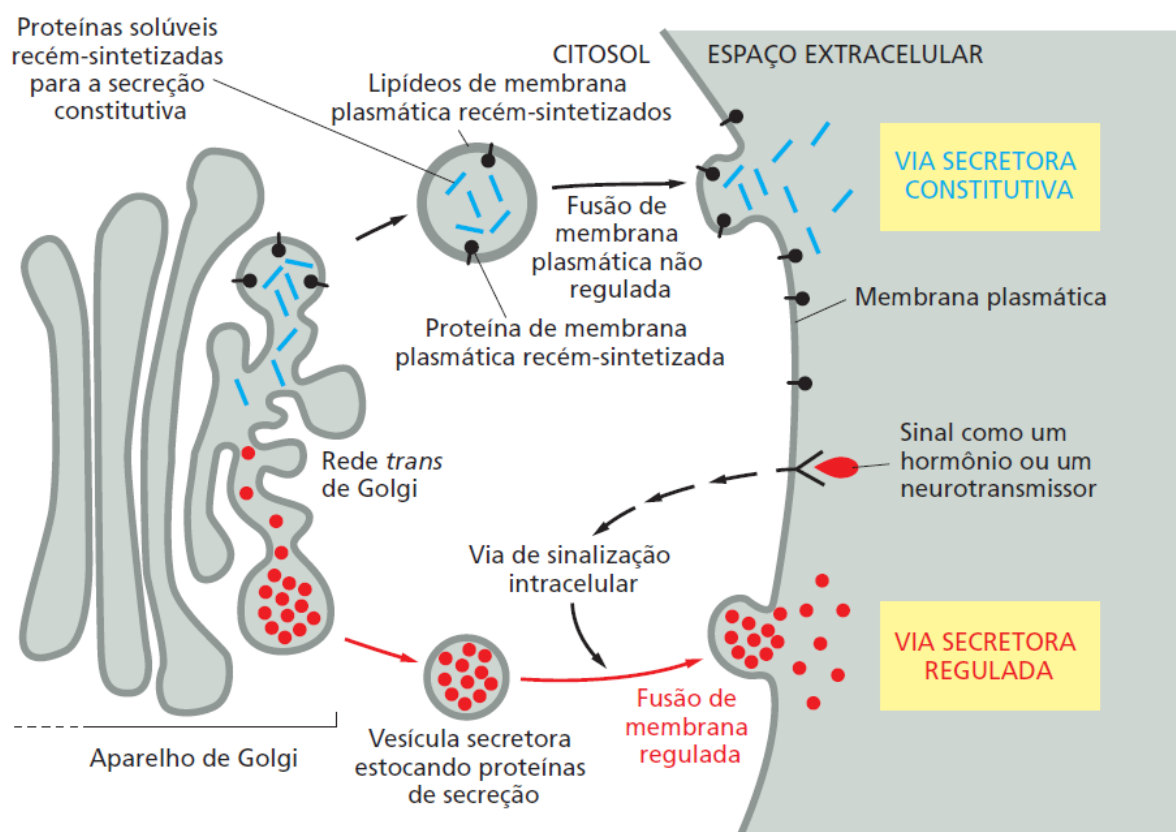


Figura 6 – Vias secretoras constitutiva e regulada

Fonte: Extraído e adaptado ALBERTS *et al.*, 2017, p. 742

As proteínas destinadas às vesículas secretoras, chamadas de proteínas de secreção ou secretoras, são empacotadas em vesículas apropriadas na TGN através de um mecanismo que envolve um agrupamento seletivo de proteínas de secreção. Os sinais que direcionam essas proteínas de secreção não são bem definidos e podem ser muito diversos.

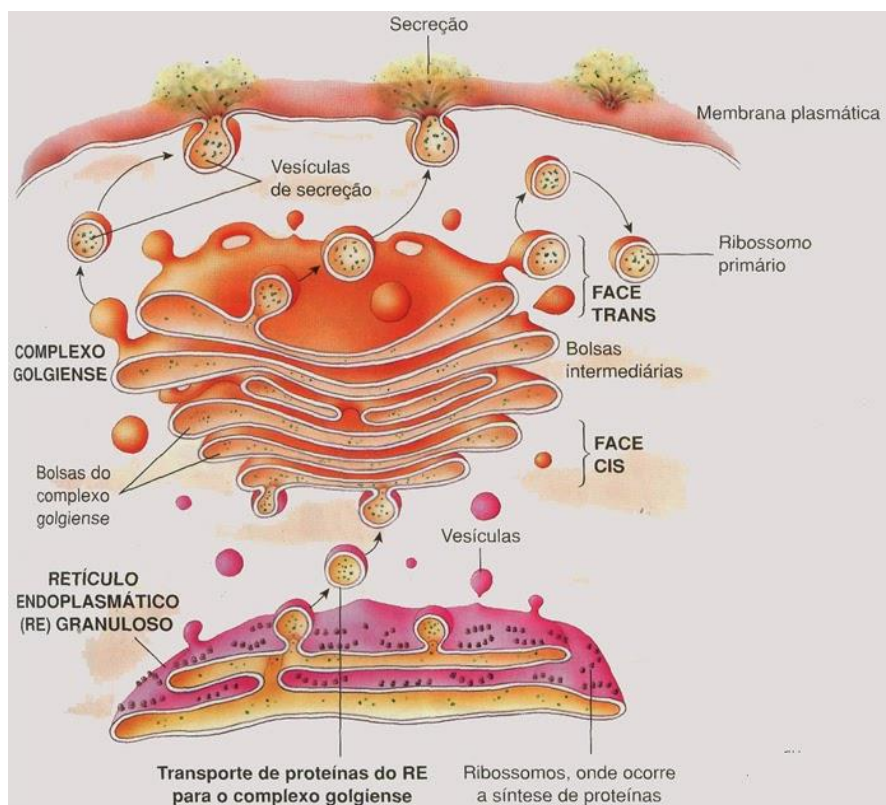
No início, estas vesículas encontram-se imaturas e se parecem com cisternas dilatadas da rede *trans* de Golgi que saíram do complexo de Golgi; contudo, conforme as vesículas amadurecem, elas se fundem umas às outras e se tornam concentradas de proteínas. Após carregadas, essas vesículas secretoras chegam ao local de secreção, muitas vezes distante da TGN, e esperam próximas da membrana plasmática até que recebam um sinal para que secretem seu conteúdo para então se fundirem na membrana (ALBERTS, B.; BRAY, D.; LEWIS, J.; RAFF, M.; ROBERTS, K.; WATSON, 2017; BURGESS; KELLY, 1987).

Os exossomos são um subconjunto de vesículas extracelulares rodeado por uma membrana de bicamada lipídica e secretado pela maioria células eucarióticas (RUIVO *et al.*, 2017). Identificados no final da década de 1980, os exossomos eram originalmente considerados como produtos celulares residuais (JOHNSTONE *et al.*, 1987); contudo, com o desenvolvimento de novas metodologias e técnicas de pesquisa, notou-se que os exossomos representam um novo modo de comunicação intercelular e contribuem para uma ampla gama de processos biológicos na saúde e na doença, incluindo câncer (KALLURI; LEBLEU, 2020).

A função biológica do exossomo depende de suas cargas bioativas, como lipídios, metabólitos, proteínas e ácidos nucleicos (PUHKA *et al.*, 2017; SCHEY; LUTHER; ROSE, 2015; SKOTLAND; SANDVIG; LLORENTE, 2017; VAN DEN BOORN *et al.*, 2013), que podem ser entregues às células-alvo. Crescentes evidências sugerem que exossomos derivados de tumor desempenham papéis críticos no câncer. Portanto, exossomos e suas cargas podem servir como marcadores de prognóstico do câncer, alvos ou mesmo como portadores de drogas anticâncer (TAI *et al.*, 2018).

A Figura 7 sintetiza a via de secreção proteica convencional do RE até ser secretada.

Figura 7 – Visão geral da via de secreção proteica convencional



Fonte: Imagem extraída do site: <<http://mayelizi.blogspot.com/2015/04/o-que-e-complexo-de-golgi-o-complexo-de.html>>

2.4 Vias de secreção proteica não convencionais

A maioria das proteínas secretadas seguem a via de secreção proteica convencional, que envolve um caminho bem documentado das proteínas do RE ao complexo de Golgi, até serem finalmente secretadas para o exterior da célula. No entanto, não é incomum o uso de vias alternativas de secreção proteica, que não sigam a ordem RE–Golgi. Recentemente, tem sido mostrado que um número significativo de proteínas alcança a membrana plasmática ou o espaço extracelular por diferentes vias. Estas vias são chamadas de vias de secreção proteica não convencionais (UPS; do inglês, *unconventional protein secretion*) (RABOUILLE, 2017).

Estas vias não convencionais podem ser divididas em dois tipos: (i) secreção extracelular de proteínas citosólicas que não carregam o peptídeo sinal, por exemplo, as proteínas *leaderless* (proteínas sem líder, em uma tradução livre); (ii) tráfego na superfície celular de proteínas transmembrana, contendo peptídeo-sinal, através de uma via que ignora o Golgi.

As *proteínas learderless*, por sua vez, podem ser secretadas através da mediação de três mecanismos: Tipo I, ou translocação direta da proteína mediada por poros através da membrana plasmática; Tipo II, secreção baseada em transportadores ABC (comum em leveduras); e o Tipo III ou secreção baseada em autofagossomo/endossomo.

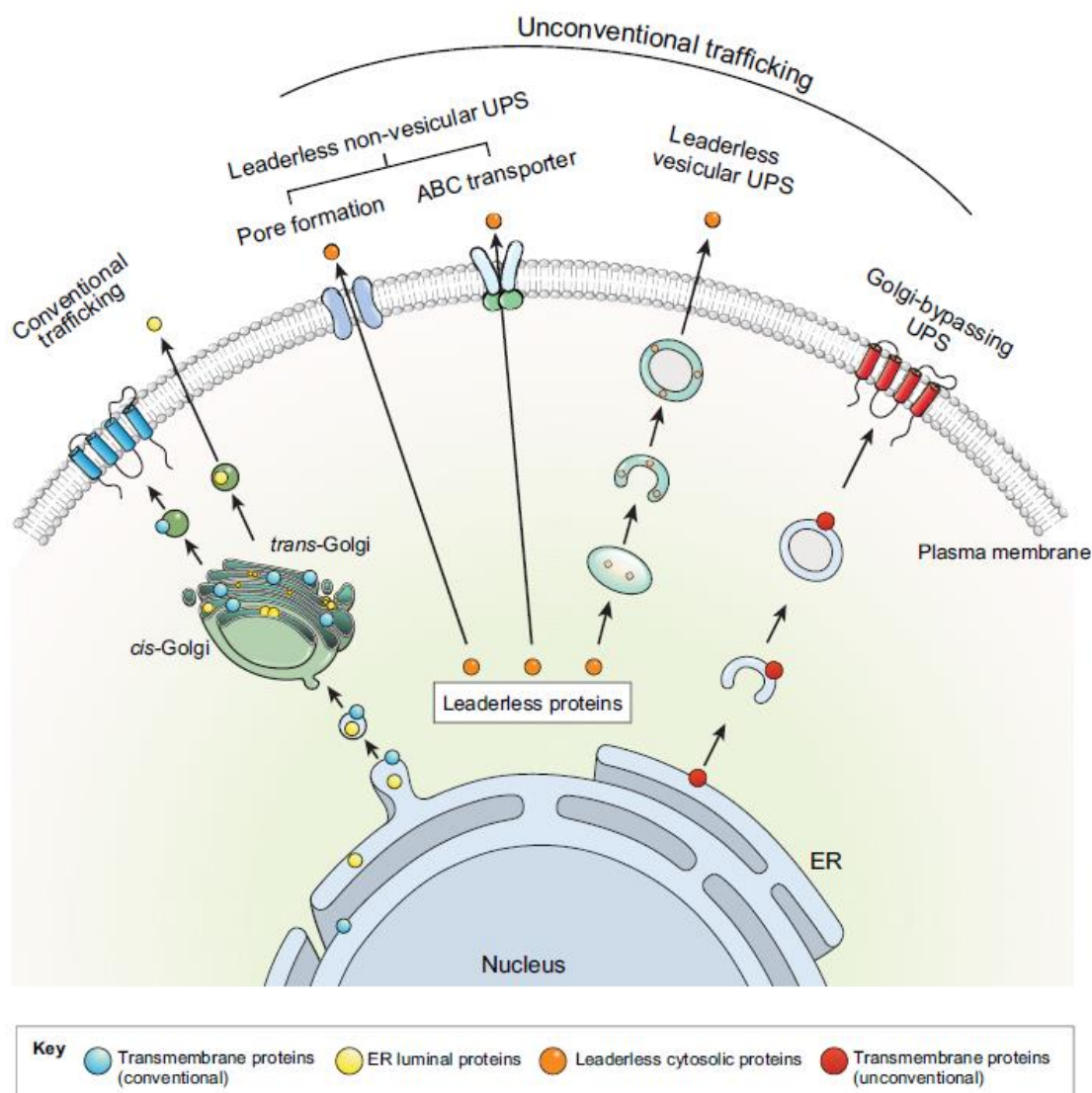
A segunda categoria (tráfego na superfície celular de proteínas transmembrana) é composta por proteínas com peptídeo sinal e/ou domínio transmembrana que entram no RE mas desviam seu caminho do complexo de Golgi e vão direto à membrana plasmática, este também é chamado de tipo IV ou *Golgi-bypass* (Figura 8) (KIM; GEE; LEE, 2018; RABOUILLE, 2017).

A maioria das vias UPS são induzidas por diferentes tipos de estresse à célula, tais como falta de nutrientes (CRUZ-GARCIA *et al.*, 2014), estresse mecânico (SCHOTMAN; KARHINEN; RABOUILLE, 2008), inflamação (SCHRODER; TSCHOPP, 2010) e estresse no RE (GEE *et al.*, 2011). Observa-se que muitas doenças estão associadas com diversos tipos de estresse não só no nível celular, mas também no nível do organismo (FULDA *et al.*, 2010).

Estudos recentes têm mostrado que *heat shock proteins* (HSPs) são secretadas de forma não convencional e participam de um papel essencial na imunomodulação, proliferação, invasão e angiogênese em processos tumorais (RODRÍGUEZ *et al.*, 2009; SANTOS; MARTINS; HAJJ, 2017). Por exemplo, a proteína HSPH1, quando secretada por células de câncer de cólon, induz a diferenciação de macrófagos, favorecendo um perfil pró-tumoral e anti-inflamatório (BERTHENET *et al.*, 2016); além disso, HSPB1 secretada por células de câncer de mama, ativa a diferenciação de monócitos para macrófagos com fenótipos imunotolerantes que perdem a atividade tumoricida e se tornam pró-angiogênicos (ARISPE; DOH; DE MAIO, 2002).

Ademais, há o registro de proteínas intracelulares que possuem funções diferentes quando secretadas, contribuindo para invasão de tecidos, o que sugere, inclusive, a presença de novas vias de secreção proteica não convencionais que ainda não foram completamente elucidadas (MÉNDEZ *et al.*, 2019; SCAFFIDI; MISTELI; BIANCHI, 2002). Deste modo, o estudo das UPS indica um grande potencial para o melhor entendimento e desenvolvimento de tratamentos para diversas doenças, sobretudo o câncer.

Figura 8 – Visão geral das vias de secreção proteica convencional e não convencionais



Fonte: Extraído e adaptado de KIM; GEE; LEE, 2018

Nota: Visão geral das vias de secreção proteica tanto convencional quanto não convencionais. Observa-se à esquerda a via de secreção convencional e à direita as quatro vias de secreção não convencionais conhecidas. Os tipos de proteínas secretadas estão divididos em um código de cores: proteínas transmembrana (via convencional) em azul, proteínas do RE (amarelo), proteínas *leaderless* citosólicas em laranja e proteínas transmembrana (via não convencional) em vermelho.

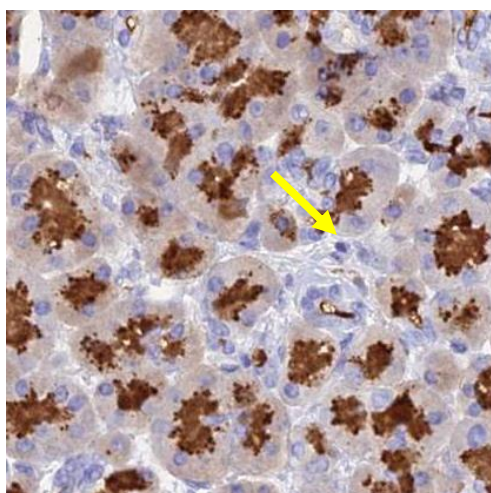
2.5 O secretoma

Todas as proteínas secretadas, assim como outras moléculas que são secretadas pelas células para o meio extracelular, formam um agrupamento definido pelo termo **secretoma**. Este termo foi primeiramente utilizado em um estudo genômico de proteínas secretadas pela bactéria *Bacillus subtilis* (TJALSMA *et al.*, 2000). Segundo a definição proposta por AGRAWAL *et al.* (2010), o secretoma constitui-se como “um grupo global de proteínas secretadas para o espaço

extracelular por uma célula, tecido, órgão, ou organismo, a qualquer momento e condição, através de mecanismos conhecidos e desconhecidos que regulam organelas secretórias”.

O secretoma também possui uma rica fonte de alvos de medicamentos e grande parte dos exames de diagnóstico de sangue utilizados na clínica são direcionados para proteínas secretadas, o que enfatiza a importância dessa classe de proteínas para a medicina e a biologia (XUE; LU; LAI, 2008). Ademais, as proteínas secretadas mais abundantes incluem enzimas pancreáticas como a *Chymotrypsin-like elastase family, member 3A* (CELA3A; Figura 9) e outras enzimas digestivas expressas na glândula salivar ou estômago (*The Human Protein Atlas*; <https://www.proteinatlas.org/>; THUL *et al.*, 2017; UHLEN *et al.*, 2017).

Figura 9 – CELA3 em amostra histológica no pâncreas



50 μ m

Fonte: *The Protein Atlas*

Nota. Amostra histológica mostrando a localização (seta amarela) da proteína secretada CELA3A no pâncreas.

Como visto anteriormente, a sequência-sinal ou peptídeo sinal indica para qual compartimento a proteína será encaminhada, inclusive se será encaminhada para fora da célula. Um peptídeo sinal possui uma série de recursos que permitem sua previsão computacional com base na sequência de aminoácidos da proteína (YUAN *et al.*, 2003).

No banco de dados *The Human Protein Atlas*, o secretoma humano foi predito por uma varredura do proteoma inteiro usando três métodos para predição de peptídeo-sinal: SignalP4.0 (<http://www.cbs.dtu.dk/services/SignalP/>; ALMAGRO ARMENTEROS *et al.*, 2019), Phobius (<https://phobius.sbc.su.se/>; (KÄLL; KROGH; SONNHAMMER, 2007) e SPOCTOPUS (<http://octopus.cbr.su.se/>; (VIKLUND *et al.*, 2008), os quais mostraram dar resultados de predição confiáveis em uma análise comparativa. Todas as proteínas com um peptídeo sinal previsto por pelo menos dois dos três métodos foram consideradas secretadas e foram posteriormente anotadas manualmente a fim de excluir genes que são previstos por residir em locais intracelulares, como RE ou complexo de Golgi, apesar de terem uma previsão de peptídeo sinal (THUL *et al.*, 2017).

Os números resultantes de proteínas secretadas previstas são mostrados na Tabela 1.

Tabela 1 – Predição do secretoma humano por quatro métodos de predição diferentes

Método utilizado	Número de proteínas	Referência
Proteínas secretadas previstas em <i>The Human Protein Atlas</i>	4361	(THUL <i>et al.</i> , 2017; UHLEN <i>et al.</i> , 2017; UHLÉN <i>et al.</i> , 2015)
Proteínas secretadas previstas em SignalP	5816	(ALMAGRO ARMENTEROS <i>et al.</i> , 2019)
Proteínas secretadas previstas em Phobius	7613	(KÄLL; KROGH; SONNHAMMER, 2007)
Proteínas secretadas previstas em SPOCTOPUS	8165	(VIKLUND <i>et al.</i> , 2008)

Nota. O número de proteínas secretadas em *The Human Protein Atlas* é consideravelmente menor pois na metodologia utilizada foram excluídas proteínas intracelulares que continham uma previsão de peptídeo sinal.

2.6 Secretoma e câncer

As funções biológicas das proteínas secretadas são diversas, compreendendo desde moléculas sinalizadoras que participam de processos celulares como diferenciação, apoptose e angiogênese, até componentes envolvidos em processos patológicos, como metástase, em diversos tipos de câncer (PATEL *et al.*, 2014).

Processos patológicos como o câncer têm um importante reflexo na síntese proteica e, conseqüentemente, no repertório de proteínas secretadas por células tumorais em relação às células não tumorais (PALTRIDGE; BELLE; KHEW-

GOODALL, 2013); deste modo, o secretoma de células tumorais pode ser uma fonte promissora de possíveis biomarcadores, pois proteínas secretadas também podem ser detectadas em fluídos biológicos como saliva, sangue e urina, comumente analisados em exames de rotina e/ou de acompanhamento do prognóstico de pacientes (LIN *et al.*, 2013).

As células tumorais frequentemente recrutam células do estroma tumoral para contribuir no processo tumorigênico. Esta contribuição pode ocorrer tanto pela indução na produção e liberação de sinais moleculares, que atuem na progressão tumoral, quanto pelo auxílio na dispersão de células tumorais, por meio da atividade de proteases que degradam a matriz extracelular. (HANAHAAN; WEINBERG, 2000, 2011; KESSENBROCK; PLAKS; WERB, 2010).

Neste processo, a secreção de moléculas específicas feita por células tumorais torna-se uma estratégia de grande eficácia para mediação dos eventos de sinalização celular, contribuindo direta ou indiretamente para o desenvolvimento e proliferação das células cancerosas (HANAHAAN; WEINBERG, 2000, 2011).

Abordagens experimentais sistêmicas voltadas para a caracterização de secretomas celulares fornecem indícios qualitativos e quantitativos de extrema importância para o entendimento da biologia do processo tumoral (SCHAAIJ-VISSER *et al.*, 2013). No contexto da biologia tumoral, as proteínas secretadas possuem um papel fundamental na comunicação intercelular e no desenvolvimento e progressão de tumores (CACCIA *et al.*, 2011; LIN *et al.*, 2013; PALTRIDGE; BELLE; KHEW-GOODALL, 2013); deste modo, o estudo de secretomas tumorais tem propiciado uma visão sistêmica e funcional do repertório de moléculas potencialmente envolvidas nos eventos bioquímicos associados à ascensão tumoral em tecidos e órgãos. Sendo assim, atualmente, linhas de pesquisa em oncologia têm explorado cada vez mais o uso de secretomas para o entendimento dos processos de comunicação celular e interação das células tumorais com o microambiente tumoral (BROWN *et al.*, 2012; PALTRIDGE; BELLE; KHEW-GOODALL, 2013).

2.7 SUMOilação

Modificações pós-traducionais (MPTs) emergiram como mecanismos regulatórios moleculares cruciais capazes de controlar a especificidade, tempo, duração e amplitude de todos os processos fisiológicos (FLOTHO; MELCHIOR, 2013).

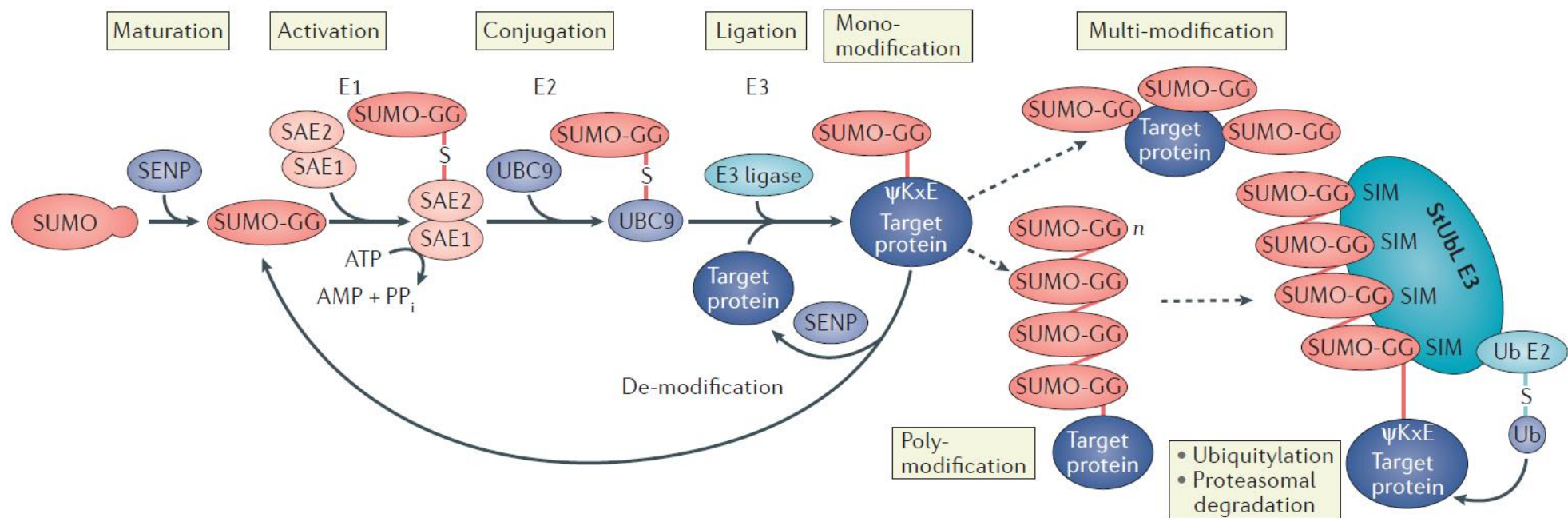
As proteínas modificadoras do tipo ubiquitina (**SUMO**; do inglês, **S**mall **U**biquitin-like **M**odifier) são uma família de pequenas proteínas que são covalentemente ligadas e separadas de outras proteínas nas células para modificar sua função, este processo é chamado de SUMOilação (HAY, 2005).

A SUMOilação é uma MPT que ocorre principalmente na porção lisina das proteínas (GEISS-FRIEDLANDER; MELCHIOR, 2007) e desempenha um papel importante na manutenção e estabilidade proteica, melhorando a capacidade de resposta ao estresse das células (ENSERINK, 2015; NISKANEN; PALVIMO, 2017; WILKINSON; HENLEY, 2010). Ela foi relatada pela primeira vez na década de 1990 (MAHAJAN et al., 1997; MATUNIS; COUTAVAS; BLOBEL, 1996); desde então, suas funções têm sido estudadas em diversas doenças regulando a expressão e função de diferentes proteínas (GAREAU; LIMA, 2010). Há quatro membros da família SUMO nos mamíferos: SUMO-1, SUMO-2, SUMO-3 e SUMO-4 (ENSERINK, 2015; JOHNSON, 2004; KUREPA et al., 2003), os quais são altamente conservados em todos os eucariotos (YANG et al., 2017).

A SUMOilação também pode ser revertida (deSUMOilação) por proteases SUMO. O equilíbrio entre a conjugação SUMO e a desconjugação exerce uma profunda influência na função fisiológica da célula. As duas funções das proteases SUMO incluem tanto a maturação dos membros da família SUMO quanto a remoção de SUMO das proteínas alvo (DRAG; SALVESEN, 2008; YEH, 2009). Semelhante à ubiquitina e muitas outras proteínas semelhantes à ubiquitina (Proteínas Ubl), todos os membros da família SUMO precisam em primeiro lugar uma protease para transformar os precursores imaturos na forma madura. As proteínas subunidade 1 da enzima de ativação do SUMO (SAE1; do inglês, SUMO-activating enzyme subunit 1) e enzima de ativação 1 do tipo Ubiquitina E1B (UBA2 ou SAE2; do inglês, Ubiquitin-like 1-activating enzyme E1B) formam um heterodímero que funciona como uma enzima de ativação da SUMO para a SUMOilação de proteínas (KWON; CIECHANOVER, 2017).

A Figura 10 elucida a via da SUMOilação.

Figura 10 – A via SUMO



Fonte: Extraído e adaptado de SEELER, 2017.

Nota. As proteínas modificadoras do tipo ubiquitina (SUMO) são sintetizadas como pré-proteínas que são primeiro clivadas por uma peptidase específica de SUMO1/sentrina para expor a diglicina (GG) do C-terminal (maturação). A etapa de ativação requer ATP para a ativação de SUMO1 (SAE1/SAE2) E1. A ativação da enzima então produz um tioéster SUMO-SAE2. SUMO é então transferido para a enzima de conjugação E2 UBC9, formando novamente um tioéster, antes de ser ligado, através de uma ligação isopeptídica, a um resíduo de lisina da proteína alvo.

Embora as funções específicas dependentes de SUMO para a grande maioria das proteínas permaneçam desconhecidas, muitas destas funções estão associadas a diversos processos celulares, inclusive tumorais (EIFLER; VERTEGAAL, 2015). A SUMOilação parece ter papéis diretos e indiretos não só na replicação do DNA, mas também na divisão das células e, portanto, no controle do ciclo celular (GEISS-FRIEDLANDER; MELCHIOR, 2007; JOHNSON, 2004). Ademais, o envolvimento da SUMOilação em mecanismos de manutenção do telômero também fornece uma ligação adicional com a tumorigênese (PEUSCHER; JACOBS, 2012).

3 OBJETIVOS

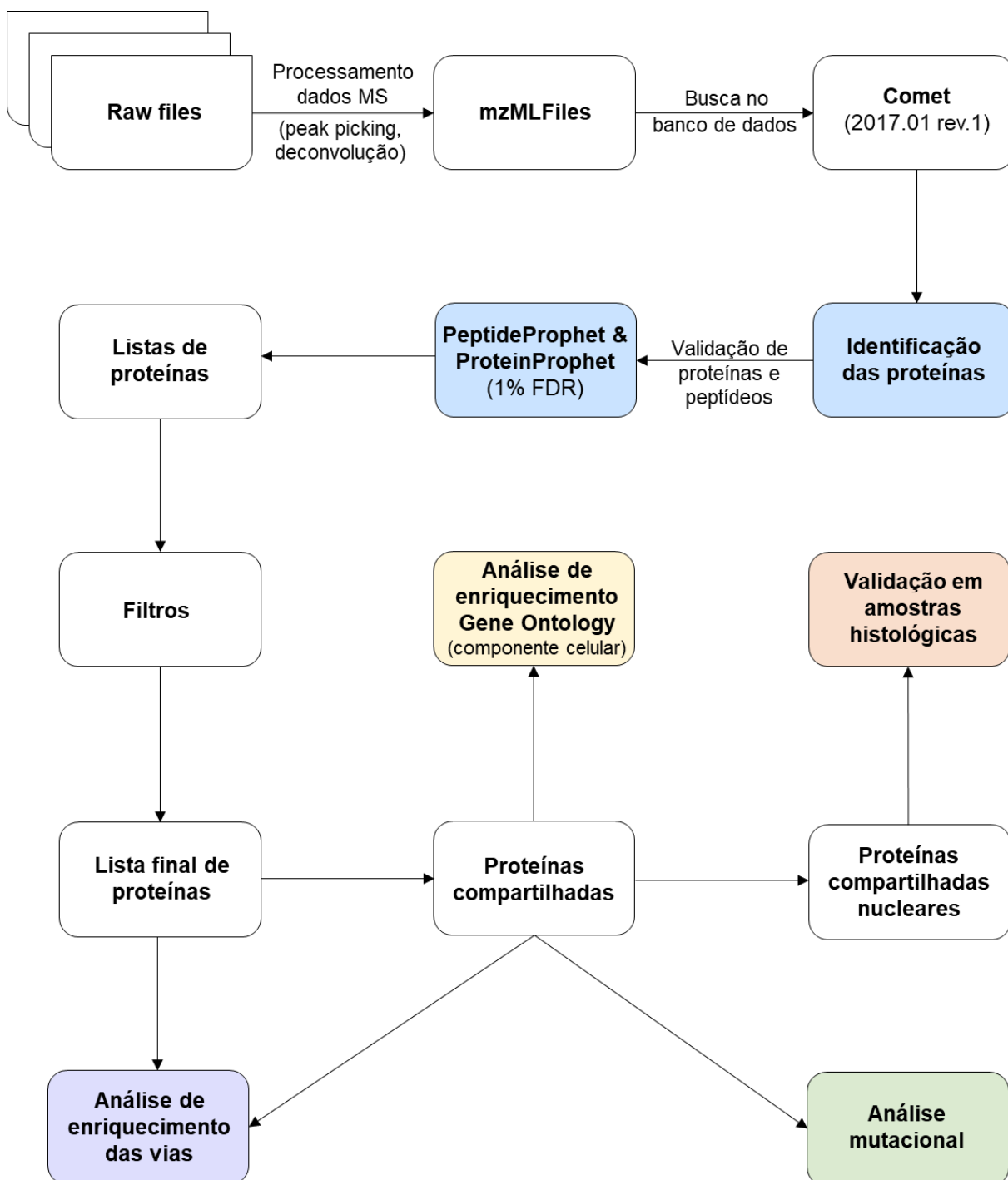
Este projeto teve como objetivo geral identificar o repertório de proteínas potencialmente secretadas de forma não canônica em diferentes secretomas tumorais que não tenham sido submetidos a nenhum tratamento experimental. Para esta finalidade, os seguintes objetivos específicos foram propostos:

1. Re-analisar corridas cromatográficas acopladas à espectrometria de massas de alta resolução (LC-MS/MS) disponíveis em repositório internacional público oriundas de análises proteômicas de secretomas de câncer de mama, melanoma, cólon e reto, ovário e sarcoma de Ewing;
2. Realizar a anotação biológica (funcional e espacial – localização celular) das proteínas encontradas após busca em banco de dados;
3. Correlacionar os resultados com as vias de secreção proteica não convencionais para identificar padrões de secreção proteica, novas funções nas proteínas, além de possíveis novas vias de secreção, utilizando *scripts* próprios e análise bioinformática.

4 MATERIAL E MÉTODOS

O fluxograma de atividades para processamento e análise dos dados está explicitado na Figura 11.

Figura 11 – Visão geral do processamento e análise dos dados avaliados nesse projeto



Fonte: O autor.

Nota. Os processos estão coloridos em **azul** para as etapas realizadas na plataforma *Trans Proteomics Pipeline* (TPP), em **roxo** para aqueles analisados pela plataforma *Reactome* (<https://reactome.org/>), em **amarelo** para aqueles analisados pelo plug-in *BiNGO*, na plataforma *Cytoscape*, em **verde** para a etapa de análise mutacional utilizando o repositório *cBioPortal* (<https://www.cbioportal.org/>) e **laranja** para a validação da localização das proteínas nucleares utilizando o banco de dados *The Human Protein Atlas* (<https://www.proteinatlas.org/>).

4.1 Aquisição dos dados de secretoma

Dados brutos de espectrômetro de massas (arquivos do tipo RAW) de secretomas tumorais de cinco diferentes tipos de câncer foram baixados do banco de dados *ProteomeXChange* (<http://www.proteomexchange.org/>; DEUTSCH *et al.*, 2017; VIZCAÍNO *et al.*, 2014).

As informações dos diferentes tipos de câncer avaliados estão descritas, em resumo, na Tabela 2. Informações adicionais sobre cada amostra biológica estão organizadas por tipo de câncer e apresentadas no Anexo A; as características técnicas dos arquivos brutos selecionados são fornecidas nos Anexos B-F.

Foram avaliados dados de amostras provenientes de sete linhagens tumorais: cólon - HCT116; mama - MDAMB231; melanoma - A375, HS895T e SH4; sarcoma de Ewing - CHLA10 e TC32, e amostras câncer de ovário de uma paciente.

Os arquivos de dados brutos relativos aos secretomas de melanoma foram obtidos no arquivo do laboratório da autora deste projeto. Os demais arquivos foram adquiridos no banco de dados do *ProteomeXChange consortium*, de acordo com a disponibilidade entre agosto e novembro de 2018.

A seleção de arquivos seguiu os seguintes critérios: (i) deveria haver completa identificação dos arquivos brutos – para obter tais informações, os respectivos autores foram contatados; (ii) os arquivos deveriam ser de secretomas tumorais provenientes de linhagens tumorais isoladas ou de amostras coletadas diretamente de humanos; sendo então, desconsiderados os secretomas de linhagens tumorais provenientes de outras espécies que não *Homo sapiens*. Além disso, os métodos de isolamento das proteínas secretadas foi similar para todos os secretomas. Estes critérios foram estabelecidos para evitar interpretações errôneas que pudessem ocorrer após os mecanismos de busca.

Tabela 2 – Arquivos selecionados no *ProteomeXChange Consortium*

Linhagem celular tumoral	ProteomeXchange ID	Número de RAW files (tamanho)	Referência
Cólon HCT116	PXD009009	44 (8.38 GB)	(CHEN <i>et al.</i> , 2018)
Mama MDAMB231	PXD000397	4 (9.15 GB)	(COX <i>et al.</i> , 2015a, 2015b)
Melanoma A375 HS895T SH4	PXD007265	6 (4.49 GB)	(LIBERATO <i>et al.</i> , 2018)
Ovário (amostra de paciente)	PXD008047	40 (35.9 GB)	(WORZFELD <i>et al.</i> , 2018)
Sarcoma de Ewing CHLA10 TC32	PXD007909	6 (10.9 GB)	(HAWKINS <i>et al.</i> , 2018)
Tamanho total		118 (76.3 GB)	

4.2 Processamento dos arquivos brutos

Os dados brutos dos espectrômetros de massas (arquivos do tipo RAW) foram processados e tratados utilizando o software *Trans-Proteomic Pipeline* (v.5.1.0) (TPP) (Syzygy, Build 201711031215-7670) (DEUTSCH *et al.*, 2011, 2015).

Após a conversão dos arquivos brutos para o formato aberto, mzML, a busca foi realizada contra um banco de dados restrito à taxonomia *Homo sapiens* (UniProtKB, *release 11_2018*; contendo 42.424 entradas), utilizando o algoritmo de busca *Comet* (versão 2017.01 rev. 1) (ENG *et al.*, 2015; ENG; JAHAN; HOOPMANN, 2013). As buscas foram ajustadas de acordo com os parâmetros determinados por cada autor; informações sobre os parâmetros utilizados em cada busca de acordo com o tipo de câncer estão disponibilizadas nos Anexos B-F.

Os peptídeos foram validados e atribuídos às proteínas correspondentes pelo algoritmo *PeptideProphet* e *ProteinProphet*, respectivamente (KELLER *et al.*, 2002) presentes na plataforma TPP e determinado ≥ 2 peptídeos únicos/proteína para o

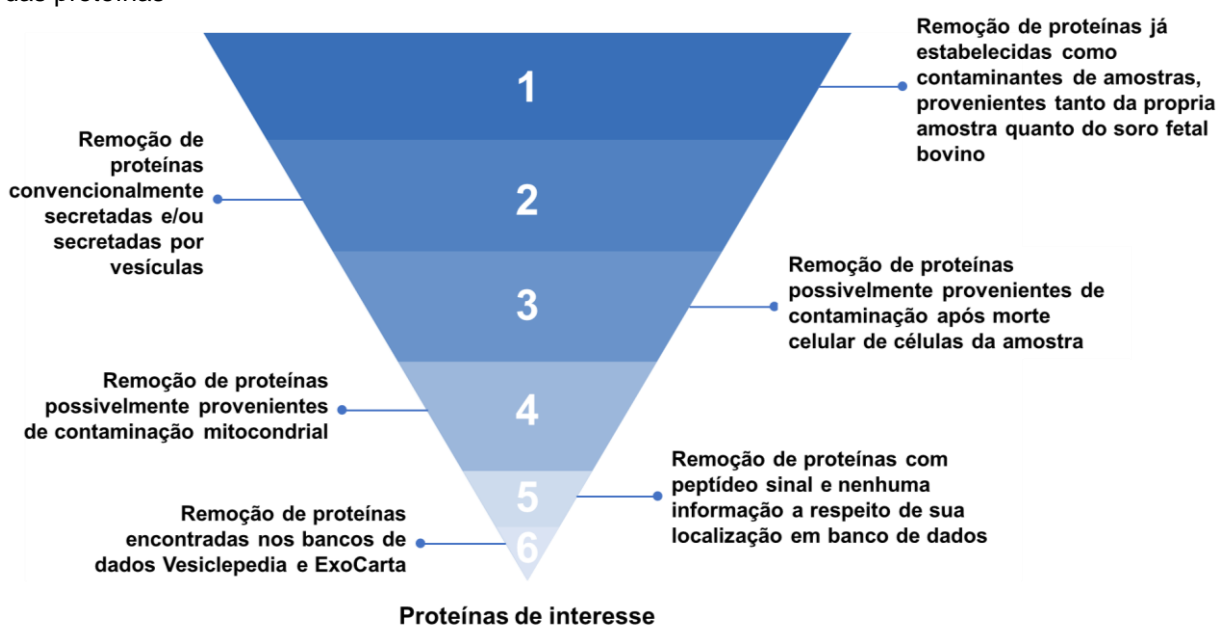
critério de identificações proteicas. As proteínas foram filtradas e exportadas em listas (uma lista de proteínas para cada linhagem celular) com a taxa de falso-positivo (FDR; do inglês, *False Discovery Rate*) definida em 1%. Logo após, as três listas de melanoma e duas de sarcoma de Ewing foram unificadas, totalizando cinco listas – uma lista para cada tipo de câncer.

4.3 Filtragem das proteínas

O processo de filtragem foi realizado nas listas de proteínas do secretoma com a seguinte estratégia: anotação manual e revisão dos dados utilizando o banco de dados UniProtKB/Swiss-Prot (BATEMAN, 2019) e a ferramenta *Gene Ontology* (GO; <http://geneontology.org>; ASHBURNER *et al.*, 2000; CARBON *et al.*, 2019).

Os filtros foram divididos em diferentes etapas através de uma busca por palavras-chave nos dados correspondentes ao GO *Cellular Component* (componente celular, em tradução livre) e à *subcellular location* (localização subcelular, em tradução livre) das proteínas, ambas categorias da ferramenta GO. A aplicação dos filtros foi dividida em cinco diferentes etapas e, para cada etapa, diferentes critérios foram estabelecidos (Figura 12).

Figura 12 – Critérios estabelecidos em cada etapa de filtragem para posterior identificação e remoção das proteínas



Fonte: O autor.

Para a identificação de proteínas cuja presença é recorrente em secretomas e decorre (provavelmente) de processo de morte celular ou aquelas advindas do soro fetal bovino (VILLARREAL *et al.*, 2013), na primeira etapa, o filtro foi aplicado para identificar e remover estas proteínas (Tabela 3).

Na segunda etapa, o filtro foi aplicado para identificar e remover proteínas que são originalmente secretadas e uma busca manual com palavras-chave (Tabela 3) foi realizada para identificá-las e removê-las; além disso, devido sua presença na via de secreção proteica convencional (ANDREI *et al.*, 1999; MACKENZIE *et al.*, 2001; QU *et al.*, 2007), proteínas que correspondem à secreção por vesículas também foram identificadas e removidas.

Na terceira etapa, um novo filtro foi aplicado para evitar a presença de possíveis contaminantes após morte celular de células das amostras; com isso, todas as proteínas que são relacionadas à apoptose em algum nível ou estão presentes na membrana celular foram manualmente identificadas e removidas.

Uma vez que a presença de proteínas mitocondriais foi verificada em alguns secretomas, optamos por remover estas proteínas, bem como proteínas da via apoptótica (TAIT; GREEN, 2010). As listas de proteínas removidas, tanto mitocondriais quanto relacionadas à apoptose, podem ser encontradas no Anexo G.

Na quarta etapa, todas as proteínas mitocondriais identificadas nos secretomas foram removidas. Na quinta etapa, proteínas que possuíam peptídeo sinal e/ou não tinham quaisquer informações sobre sua localização no banco de dados foram removidas. Por fim, foram selecionadas as cem proteínas estatisticamente relevantes tanto da plataforma Vesiclepedia (<http://microvesicles.org/>; PATHAN *et al.*, 2019), quanto ExoCarta (<http://www.exocarta.org/>; KEERTHIKUMAR *et al.*, 2016) para remover possíveis proteínas oriundas de vesículas e exossomos. A lista destas proteínas está disposta no Anexo H. Após o processo de filtragem, as proteínas remanescentes passaram a ser consideradas como proteínas de interesse ou proteínas possivelmente oriundas de via secretória não canônica, a fim de serem estudadas posteriormente.

Tabela 3 – Palavras-chave buscadas no banco de dados em cada etapa dos filtros para identificação e remoção das proteínas

Palavras-chave em cada etapa do Filtro

Primeira etapa	Segunda etapa	Terceira etapa	Quarta etapa	Quinta etapa	Sexta etapa
A1AT	secreted	apoptosis	mitochondria	Proteínas sem	Cem proteínas
AN32B	secretion	apoptotic	mitochondrion	peptídeo sinal e	estatisticamente
ANP32B	secretory	death	mitochondrial	que não possuíam	relevantes das
APOH	extracellular	cell membrane		qualquer	plataformas
CATD	exosome	cell surface		informação sobre	Vesiclepedia e
CTSD	vesicle	membrane		sua localização no	ExoCarta
		protein		banco de dados.	
DSG2					
FLNB					
GDIB					
HMCS1					
HMGB1					
HS90A					
HS90B					
HSP74					
K1C10					
K1C9					
K2C1					
KRT1					
KRT10					
KRT9					
MAOX					
NDKA					
NDKB					
PGM1					
SERPINA1					
SYAC					
SYEP					
SYK					
SYNC					
SYQ					
SYRC					
SYSC					
SYTC					
SYWC					
SYYC					

Nota. Todas as palavras-chave da primeira etapa consistem nos genes correspondentes às proteínas contaminantes de acordo com VILLARREAL *et al.* (2013). A busca e remoção dessas proteínas foi feita a partir da busca do nome de seu gene para evitar qualquer redundância do banco de dados.

4.4 Processamento e análise das proteínas possivelmente de via secretória não canônica

Após o processo de filtragem, as proteínas foram analisadas na plataforma Reactome (versão 75) (<https://reactome.org/>; JASSAL *et al.*, 2020) para se identificar o enriquecimento de vias biológicas no conjunto de proteínas identificadas.

4.4.1 Análise de intersecções proteicas e estudo de mutações

Ao comparar as listas de proteínas, foram identificadas proteínas em comum (ou compartilhadas) entre os secretomas dos diferentes tipos de câncer e, para melhor visualização e análise destas intersecções proteicas o pacote *UpSetR* (CONWAY; LEX; GEHLENBORG, 2017; LEX *et al.*, 2014) foi utilizado no ambiente R. Análises de enriquecimento das proteínas foram conduzidas utilizando o software Cytoscape (versão 3.7.1) (OTASEK *et al.*, 2019; PAUL SHANNON *et al.*, 2003), com o *plugin* BINGO (MAERE; HEYMANS; KUIPER, 2005). A taxonomia *Homo sapiens* foi selecionada, assim como as categorias do *Gene Ontology* (GO): processo biológico, componente celular e função molecular; as demais configurações foram mantidas no modo padrão do programa.

Além da análise de compartimento celular, avaliamos o perfil de mutações para os genes codificadores das proteínas compartilhadas através da plataforma *cBioPortal for Cancer Genomics* (<https://www.cbioportal.org/>; CERAMI *et al.*, 2012; GAO *et al.*, 2013) a partir de dados tumorais de pacientes obtidos e disponibilizados pelo consórcio *The Cancer Genome Atlas (TCGA) Research Network* (<https://www.cancer.gov/tcga>), com exceção ao sarcoma de Ewing, pois não haviam dados sobre este tipo de câncer oriundo do consórcio, deste modo, foi escolhido um estudo não relacionado ao TCGA (CROMPTON *et al.*, 2014). Os demais dados foram escolhidos e coletados do projeto *TCGA PanCancer Atlas* (Tabela 4) devido à sua minuciosidade e metodologia padrão para a obtenção e disponibilidade de todos os dados (DING *et al.*, 2018; WEINSTEIN *et al.*, 2013). Por fim, a plataforma *SUMOgo* (<http://predictor.nchu.edu.tw/SUMOgo/>) foi utilizada para avaliar se as proteínas compartilhadas possuem sítios de SUMOilação.

Tabela 4 – Dados tumorais de pacientes obtidos na plataforma *cBioPortal*

Tipo de câncer	Estudo	Número de pacientes	Referência
Cólon	<i>Colorectal Adenocarcinoma</i> (TCGA, <i>PanCancer Atlas</i>)	594	(BHANDARI <i>et al.</i> , 2019;
Mama	<i>Breast Invasive Carcinoma</i> (TCGA, <i>PanCancer Atlas</i>)	1.084	ELLROTT <i>et al.</i> , 2018; GAO <i>et al.</i> , 2018;
Melanoma	<i>Skin Cutaneous Melanoma</i> (TCGA, <i>PanCancer Atlas</i>)	442	HOADLEY <i>et al.</i> , 2018; LIU <i>et al.</i> , 2018;
Ovário	<i>Ovarian Serous Cystadenocarcinoma</i> (TCGA, <i>PanCancer Atlas</i>)	585	SANCHEZ-VEGA <i>et al.</i> , 2018; TAYLOR <i>et al.</i> , 2018)
Sarcoma de Ewing	<i>Pediatric Ewing Sarcoma</i> (DFCI, <i>Cancer Discov</i> 2014)	103	(CROMPTON <i>et al.</i> , 2014)

Nota. Foi dada preferência aos estudos obtidos pelo consórcio TCGA, *PanCancer Atlas*, contudo, devido à falta de disponibilidade, outro estudo (referenciado na tabela) fora escolhido para o estudo de sarcoma de *Ewing*.

4.4.2 Validação da localização das proteínas compartilhadas nucleares através de amostras de tecido

As proteínas nucleares identificadas compartilhadas em todos os secretomas tumorais foram analisadas através de uma anotação manual no banco de dados UniProtKB/Swiss-Prot e revisão da literatura.

Ademais, com o objetivo de avaliar a localização celular destas proteínas compartilhadas nucleares, utilizamos os dados histológicos de expressão/localização celular disponíveis na plataforma *The Human Protein Atlas* (<https://www.proteinatlas.org/>; UHLÉN *et al.*, 2015) de acordo com os tecidos/tipos de câncer correspondentes, bem como para amostras de tecido normais.

5 RESULTADOS

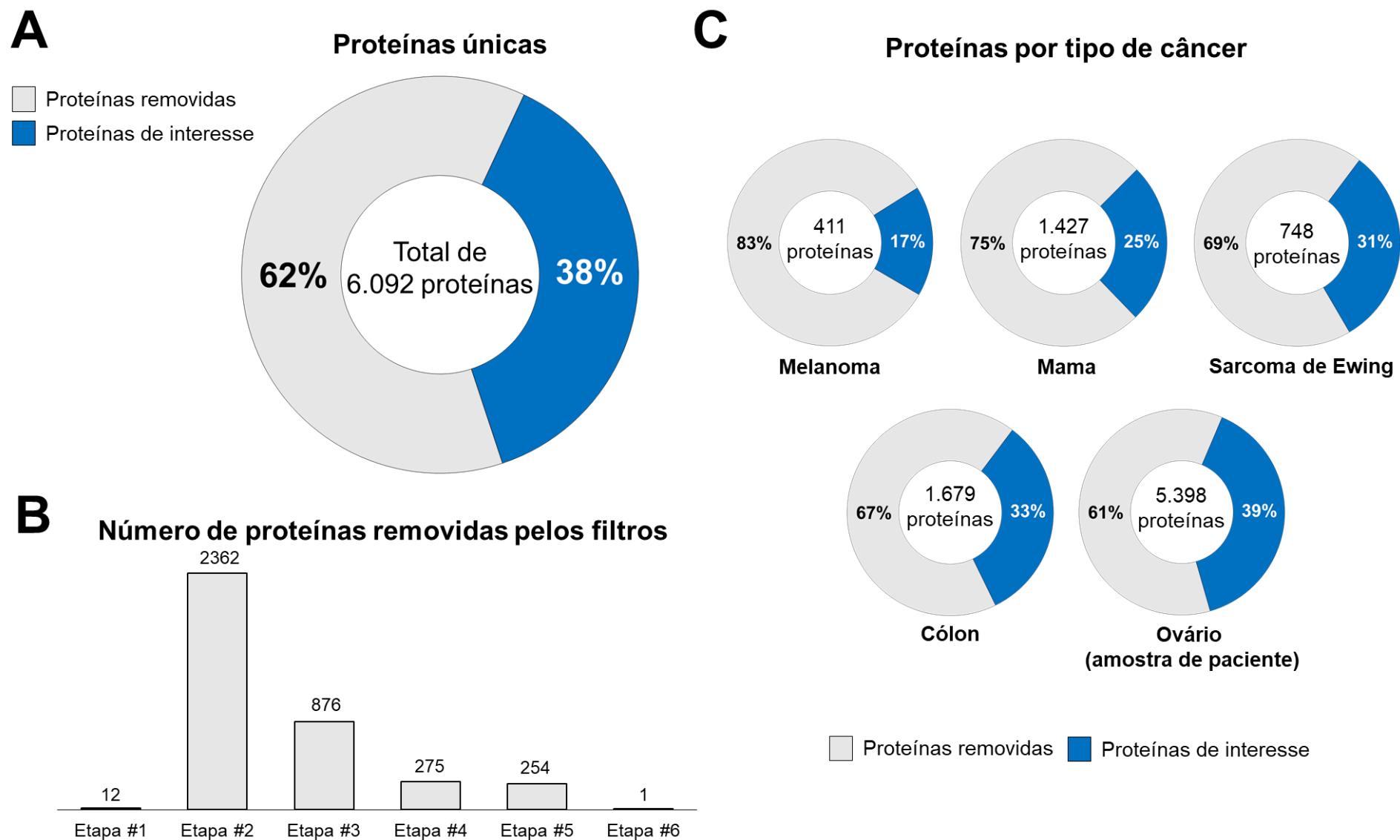
5.1 Repertório das proteínas do secretoma pós-filtragem

Para avaliar as características do repertório das proteínas presentes nos secretomas, filtros foram aplicados nas listas de proteínas geradas após exportação. Os filtros foram divididos em cinco etapas e o repertório proteico apresentado após as etapas de filtragem mostrou que 2.312 proteínas (38%), de um total de 6.092, identificadas no secretoma não eram proteínas relacionadas à secreção convencional ou derivadas de contaminação das amostras (Figura 13A). O total de proteínas removidas por cada etapa dos filtros está apresentada na Figura 13B.

Quando os secretomas são avaliados separadamente, nota-se que a porcentagem de proteínas removidas varia de 61% a 83% (Figura 13C) e a porcentagem de proteínas eliminadas em cada etapa dos filtros varia principalmente no melanoma (Figuras 13D).

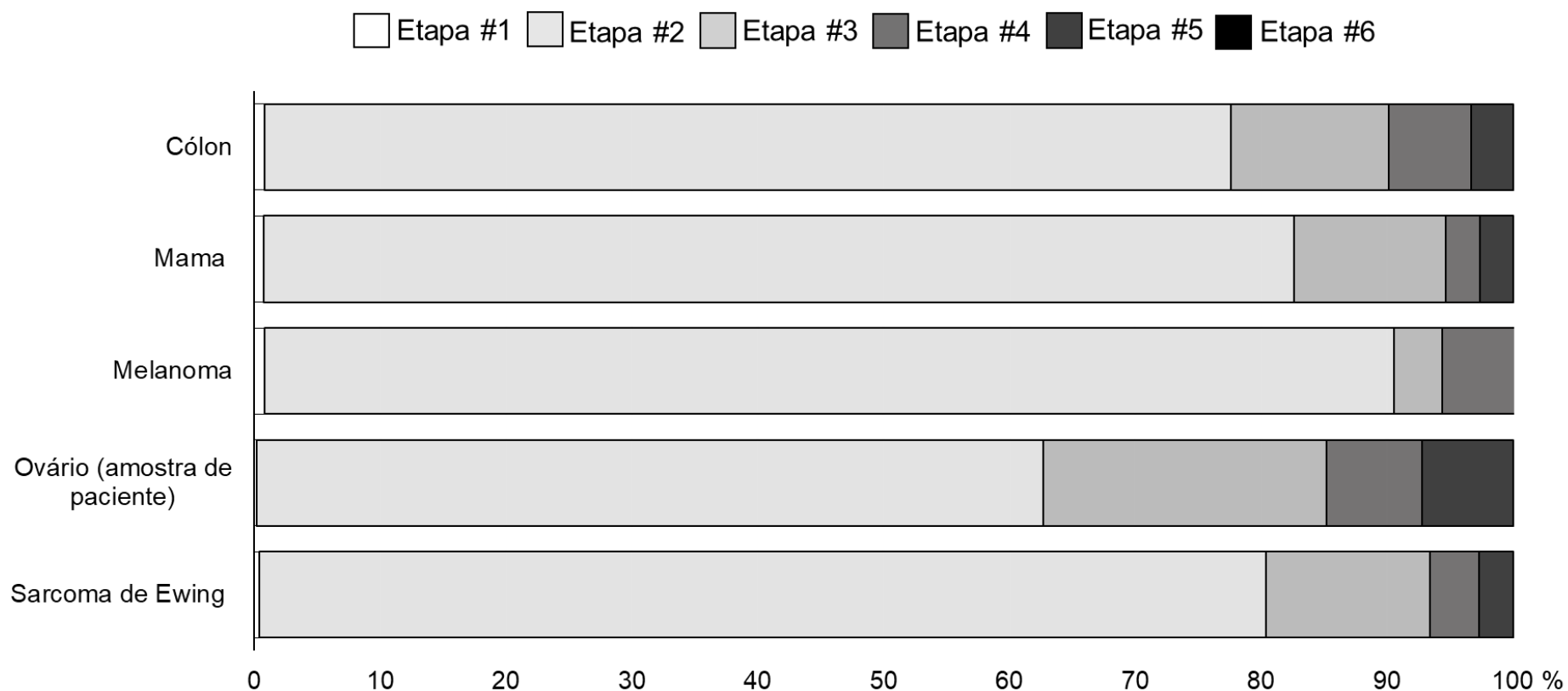
Nota-se que as proteínas possivelmente de via secretória não canônica dos secretomas tumorais em relação aos secretomas pré-filtragem variou de 17% a 39% (Figura 14A); além disso, há uma notável similaridade proteica nos secretomas pós-filtragem, sobretudo nos secretomas tumorais de cólon e ovário (Figura 14B). O número de proteínas removidas em cada etapa dos filtros – assim como todas as listas de proteínas anteriores e posteriores ao processo de filtragem – estão disponíveis na pasta *Filters*, em repositório aberto (<https://github.com/julianademorais/cancer-secretome>).

Figura 13 – Perfil geral do repertório das proteínas totais e de cada tipo de câncer



D

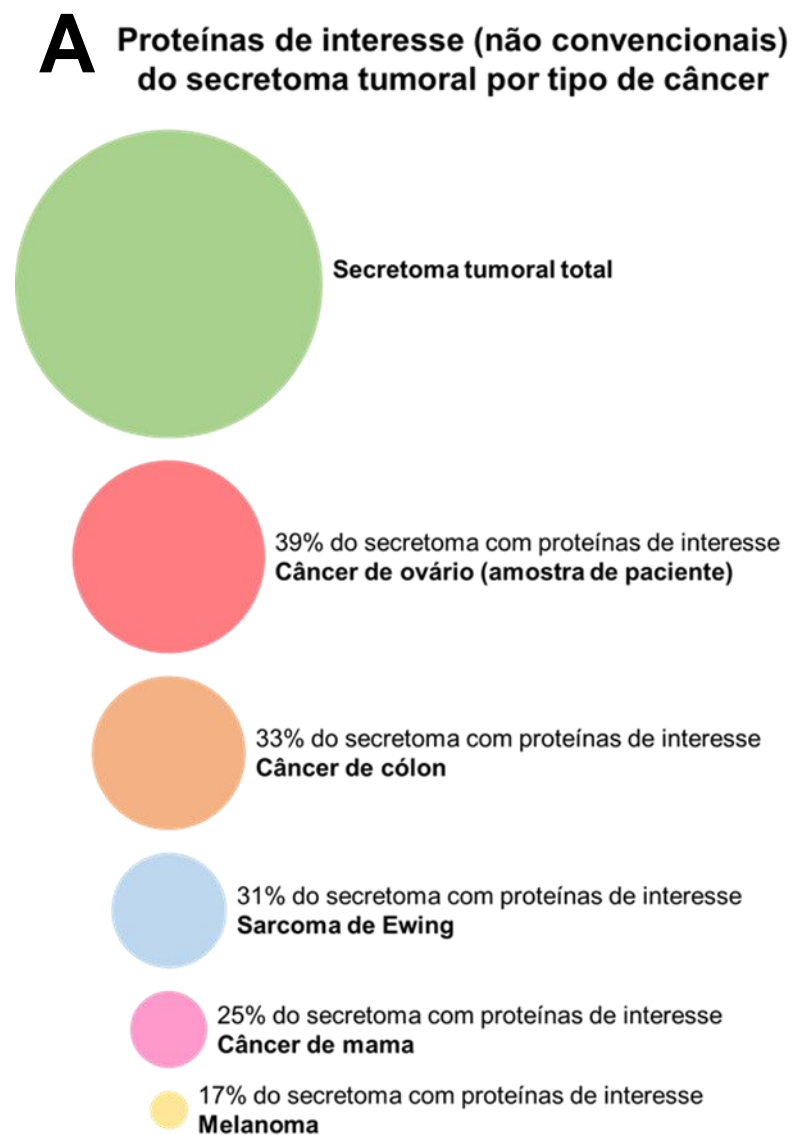
Proteínas removidas pelos filtros por etapa e tipo de câncer



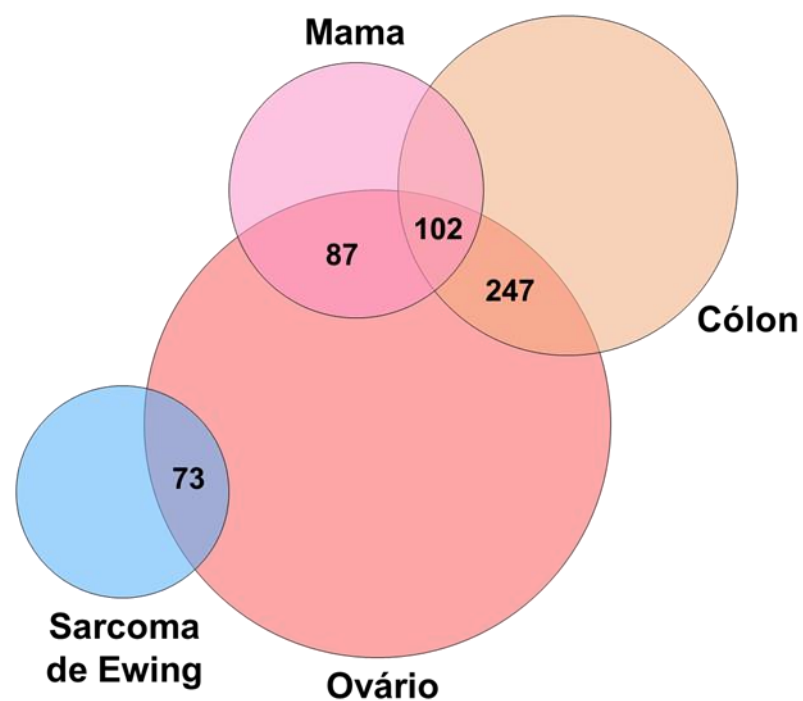
Fonte: O autor.

Nota. Análise das proteínas totais e de cada linhagem tumoral. (A) Porcentagem das proteínas totais removidas e de interesse de acordo com os filtros utilizados. (B) Número de proteínas totais removidas em cada etapa de filtragem. (C) Porcentagem das proteínas removidas e de interesse na análise dos secretomas tumorais para cada tipo de câncer. (D) Porcentagem de proteínas removidas em cada etapa de filtragem por tipo de câncer.

Figura 14 – Composição dos secretomas pós-filtragem



B Similares proteicas em diferentes secretomas tumorais



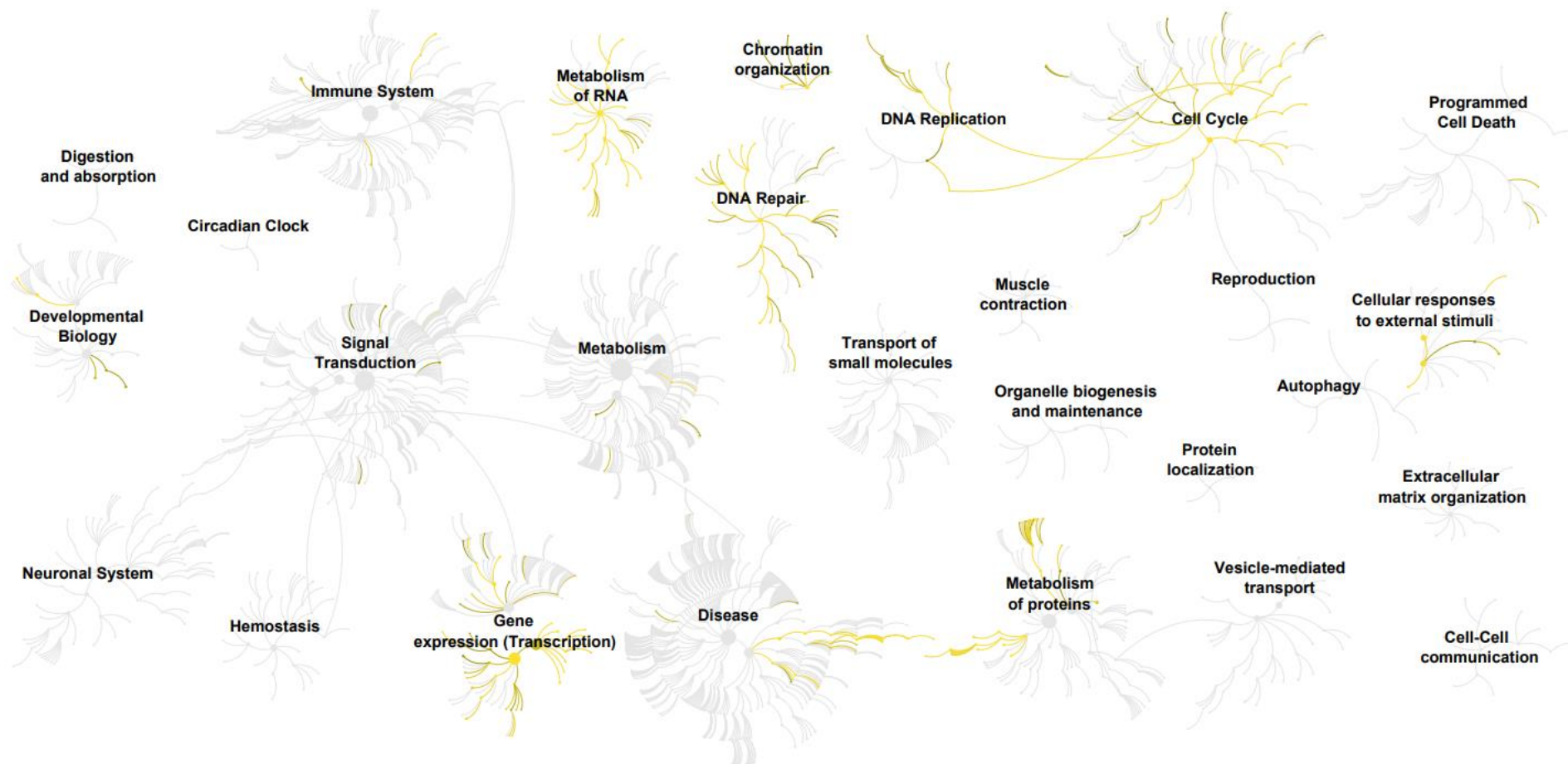
Fonte: O autor.

Nota. Composição dos secretomas pós-filtragem. (A) Porcentagem de proteínas de interesse dos secretomas tumorais em relação aos secretomas pré-filtragem. As porcentagens indicam a porção das proteínas de interesse em cada secretoma, ou seja, todas as proteínas que não foram removidas após todas as etapas de filtragem, logo, não são convencionalmente encontradas nos secretomas e são possivelmente proteínas de via secretória não canônica. (B) Similares proteicas entre os secretomas de diferentes tipos de câncer, principalmente ovário e cólon.

Quando avaliadas as vias biológicas nas quais as proteínas únicas participam, observa-se que estas estão enriquecidas principalmente nas vias do ciclo celular, metabolismo do RNA, organização da cromatina, replicação e reparo do DNA, expressão gênica e metabolismo proteico (Figura 15). A tabela completa com as vias enriquecidas mais relevantes encontra-se no Anexo I.

Quando avaliados separadamente os diferentes secretomas tumorais, nota-se que no componente celular há um enriquecimento para o núcleo (Figura 16). A Tabela 5 mostra as principais vias enriquecidas por tipo de câncer. Tanto as vias quanto o componente celular estão classificados pelo seu p-valor ajustado.

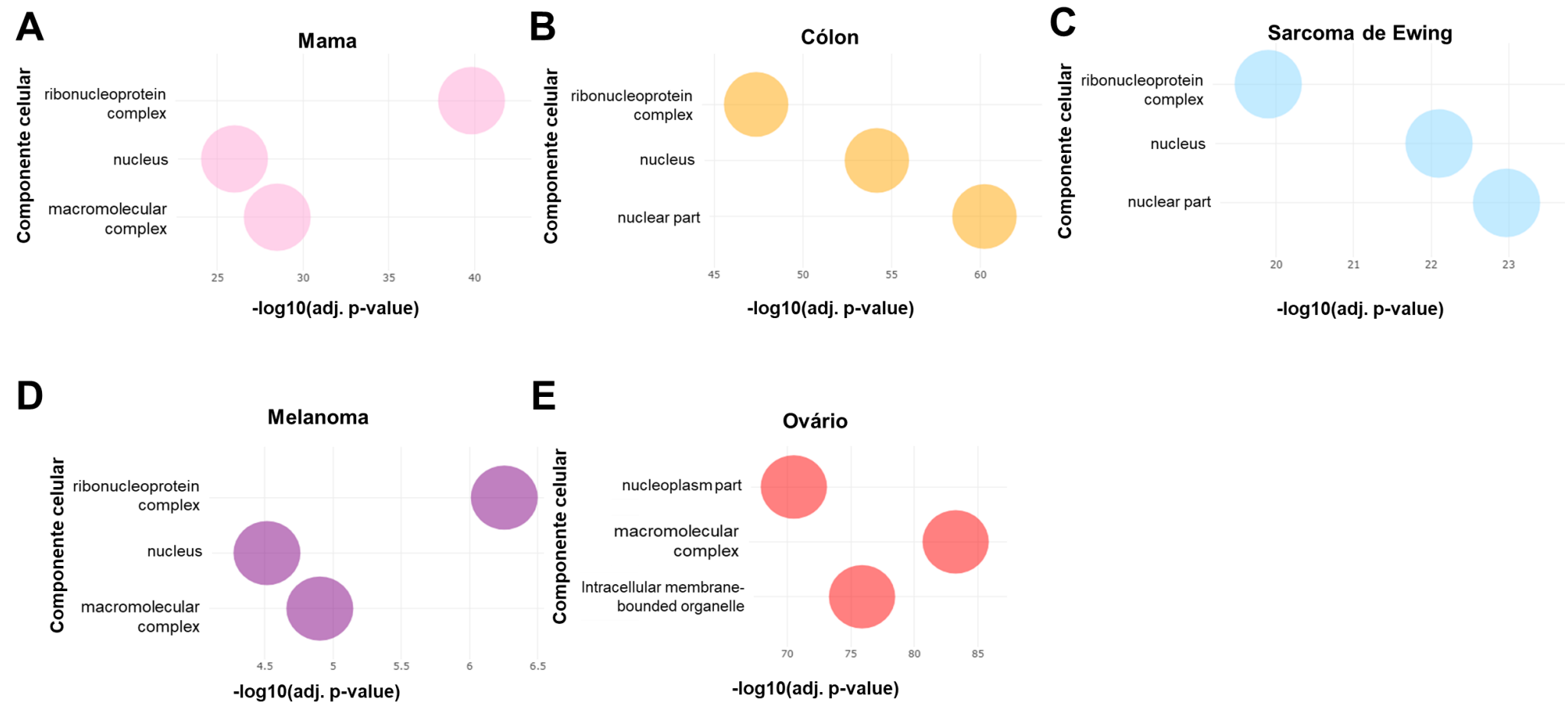
Figura 15 – Visão geral das vias enriquecidas para as proteínas únicas



Fonte: Reactome.

Nota. Análise das vias enriquecidas das proteínas únicas. As vias enriquecidas são mostradas em amarelo e as vias em cinza claro não possuem valores significantes de enriquecimento.

Figura 16 – Enriquecimento do componente celular por secretoma tumoral



Fonte: O Autor/Cytoscape.

Nota. Enriquecimento por tipo de secretoma tumoral. No componente celular, nota-se um enriquecimento para o núcleo em todos os tipos de câncer.

Tabela 5 – Principais vias enriquecidas por tipo de câncer

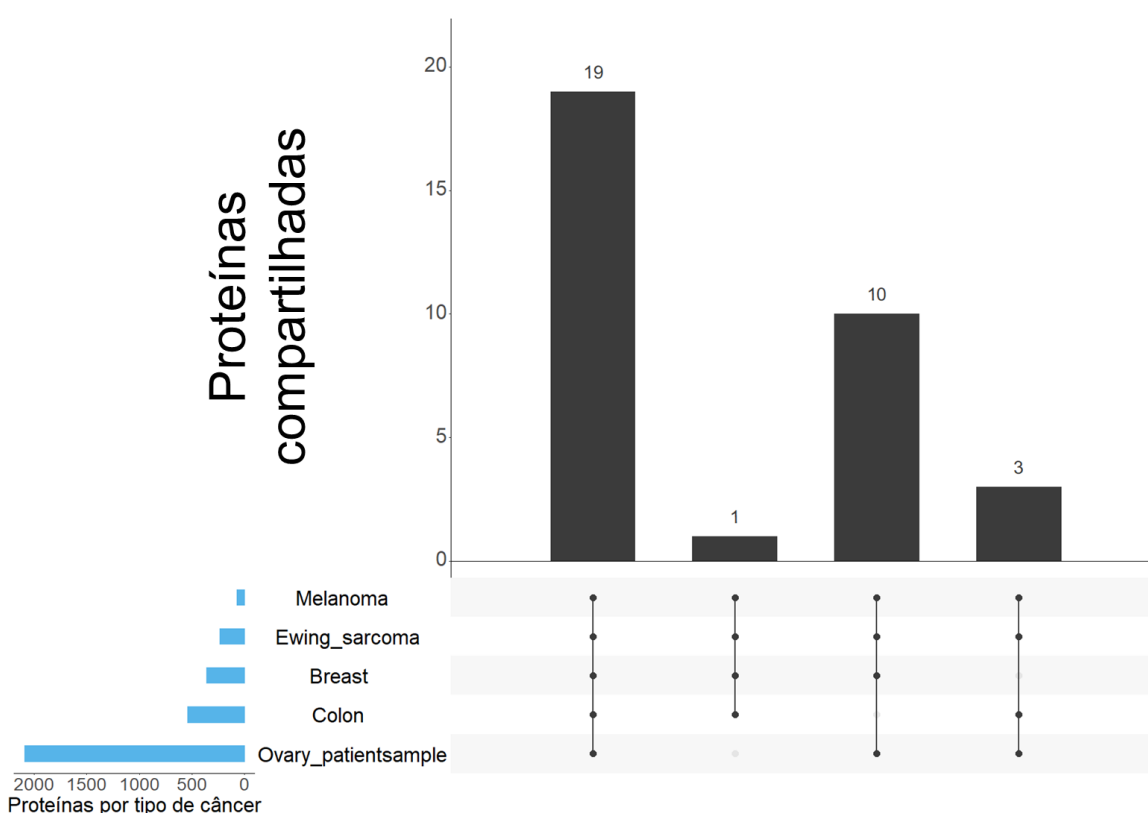
Tipo de câncer	Via biológica enriquecida	p-valor ajustado
Cólon	mRNA Splicing	1.11E-16
	GTP hydrolysis and joining of the 60S ribosomal subunit	1.11E-16
	L13a-mediated translational silencing of Ceruloplasmin expression	1.11E-16
	Eukaryotic Translation Initiation	1.11E-16
	Cap-dependent Translation Initiation	1.11E-16
Mama	Formation of a pool of free 40S subunits	1.11E-16
	GTP hydrolysis and joining of the 60S ribosomal subunit	1.11E-16
	L13a-mediated translational silencing of Ceruloplasmin expression	1.11E-16
	Nonsense-Mediated Decay (NMD)	1.11E-16
	SRP-dependent cotranslational protein targeting to membrane	1.11E-16
Melanoma	GTP hydrolysis and joining of the 60S ribosomal subunit	8.99E-05
	Eukaryotic Translation Initiation	0.00012
	SUMO is conjugated to E1 (UBA2:SAE1)	0.00049
	Translation initiation complex formation	0.00058
	Ribosomal scanning and start codon recognition	0.00058
Sarcoma de Ewing	Metabolism of RNA	4.33E-15
	GTP hydrolysis and joining of the 60S ribosomal subunit	6.18E-13
	mRNA Splicing	1.10E-12
	Cap-dependent Translation Initiation	1.81E-12
	Eukaryotic Translation Initiation	1.81E-12
Ovário	mRNA Splicing	1.11E-16
	Processing of Capped Intron-Containing Pre-mRNA	1.11E-16
	rRNA processing in the nucleus and cytosol	1.50E-13
	RNA Polymerase II Transcription Termination	6.00E-13
	mRNA 3'-end processing	4.78E-12

5.2 Análise das proteínas compartilhadas nos diferentes secretomas tumorais

5.2.1 Proteínas compartilhadas e análises de enriquecimento

Para visualizar as proteínas compartilhadas (intersecções proteicas presentes nos secretomas dos diferentes tipos de câncer) o pacote *UpSetR* foi utilizado no ambiente R. Os resultados desta análise mostram que há intersecções proteicas dentre todos os secretomas tumorais e dezenove proteínas estão presentes em cinco dos secretomas estudados (Figura 17).

Figura 17 – Visão geral das proteínas compartilhadas com maior número de intersecções entre os diferentes tipos de câncer

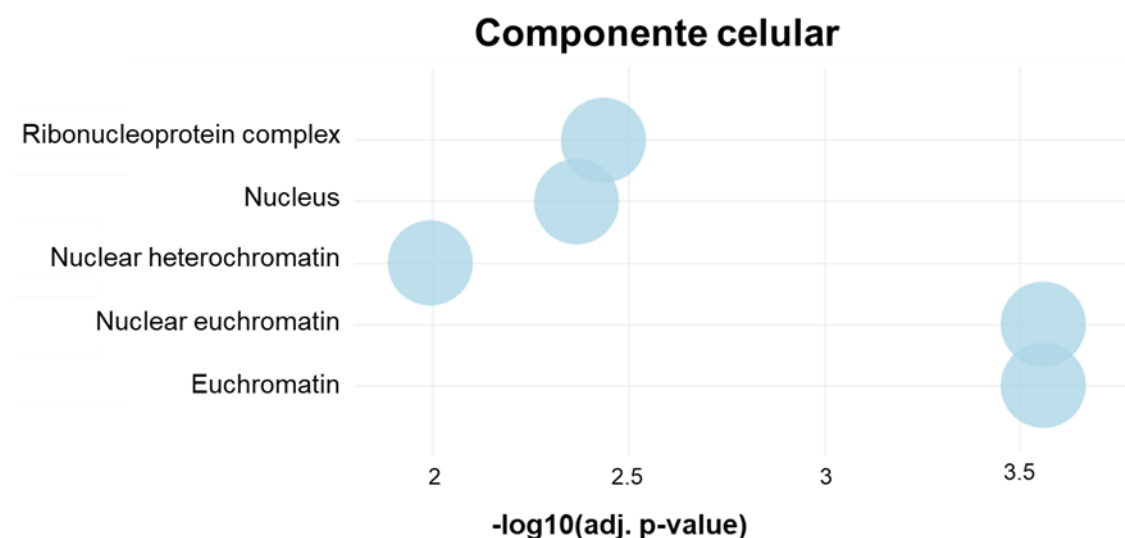


Fonte: O autor.

Nota. Visão geral das proteínas compartilhadas. Observa-se dezenove proteínas com maior número de intersecções e presentes no secretoma de todos os tipos de câncer. No gráfico, os *dots* representam a presença da proteína e o número acima de cada gráfico de barras representa quantas proteínas estão presentes no tipo de câncer.

A análise de ontologia gênica (GO; componente celular) mostrou um enriquecimento de proteínas nucleares (Figura 18).

Figura 18 – Análise de enriquecimento do componente celular das proteínas compartilhadas



Fonte: O autor.

Nota. Análise de enriquecimento do componente celular das dezenove proteínas compartilhadas. Nota-se que há um enriquecimento de proteínas nucleares.

Quando avaliadas as vias nas quais as dezenove proteínas compartilhadas estão presentes, observa-se que estas estão enriquecidas principalmente em vias de SUMOilação e processos relacionados ao envelope nuclear, como quebra do envelope e lâmina nuclear. A Tabela 6 mostra as principais vias enriquecidas, assim como as dezenove proteínas compartilhadas e seus respectivos genes. As vias estão classificadas pelo seu p-valor ajustado.

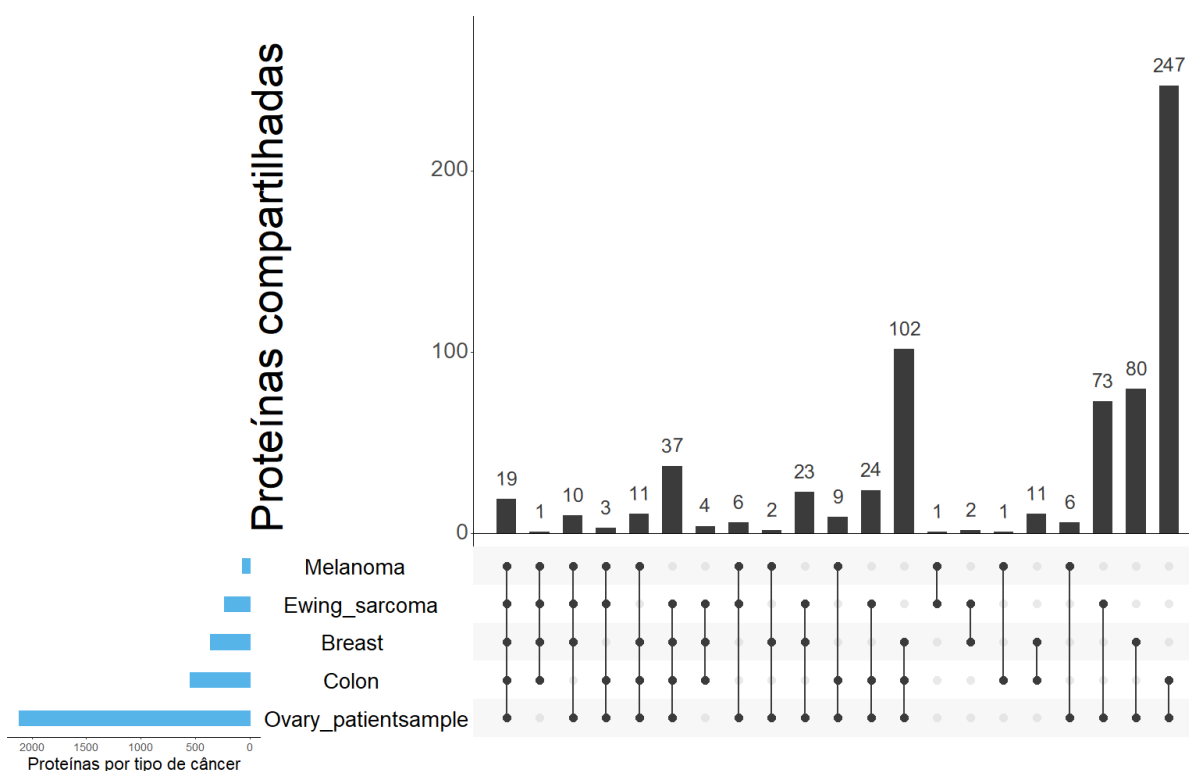
Ademais, observa-se que as dezenove proteínas compartilhadas possuem sítios de SUMOilação; os sítios de SUMOilação identificados nas proteínas compartilhadas estão dispostos no Anexo J.

Tabela 6 – Principais vias enriquecidas nas dezenove proteínas compartilhadas

Gene	ID da proteína no <i>Uniprot/Swissprot</i>	Nome da via enriquecida	p-valor ajustado
TPD52L2	O43399	SUMO is transferred from E1 to E2 (UBE2I, UBC9)	7.29E-05
TBCA	O75347		
LMNA	P02545		
SSB	P05455		
NPM1	P06748	SUMO is conjugated to E1 (UBA2:SAE1)	9.52E-05
EIF4B	P23588		
RPL13	P26373		
STIP1	P31948		
RANBP1	P43487	Processing and activation of SUMO	2.50E-04
CBX3	Q13185		
TRIM28	Q13263		
ELOC	Q15369		
SF3A1	Q15459	SUMOylation	3.44E-04
NUCKS1	Q9H1E3		
SAE1	Q9UBE0		
UBA2	Q9UBT2		
SEPTIN9	Q9UHD8	Depolymerisation of the Nuclear Lamina	7.74E-04
NUDC	Q9Y266		
STRAP	Q9Y3F4		

Além das dezenove proteínas identificadas em todos os secretomas, observa-se proteínas compartilhadas dentre diferentes secretomas tumorais. A Figura 19 ilustra estas intersecções proteicas e mostra o panorama geral das proteínas identificadas e compartilhadas em ao menos em dois tipos de câncer.

Figura 19 – Panorama geral das proteínas compartilhadas dentre os diferentes secretomas tumorais

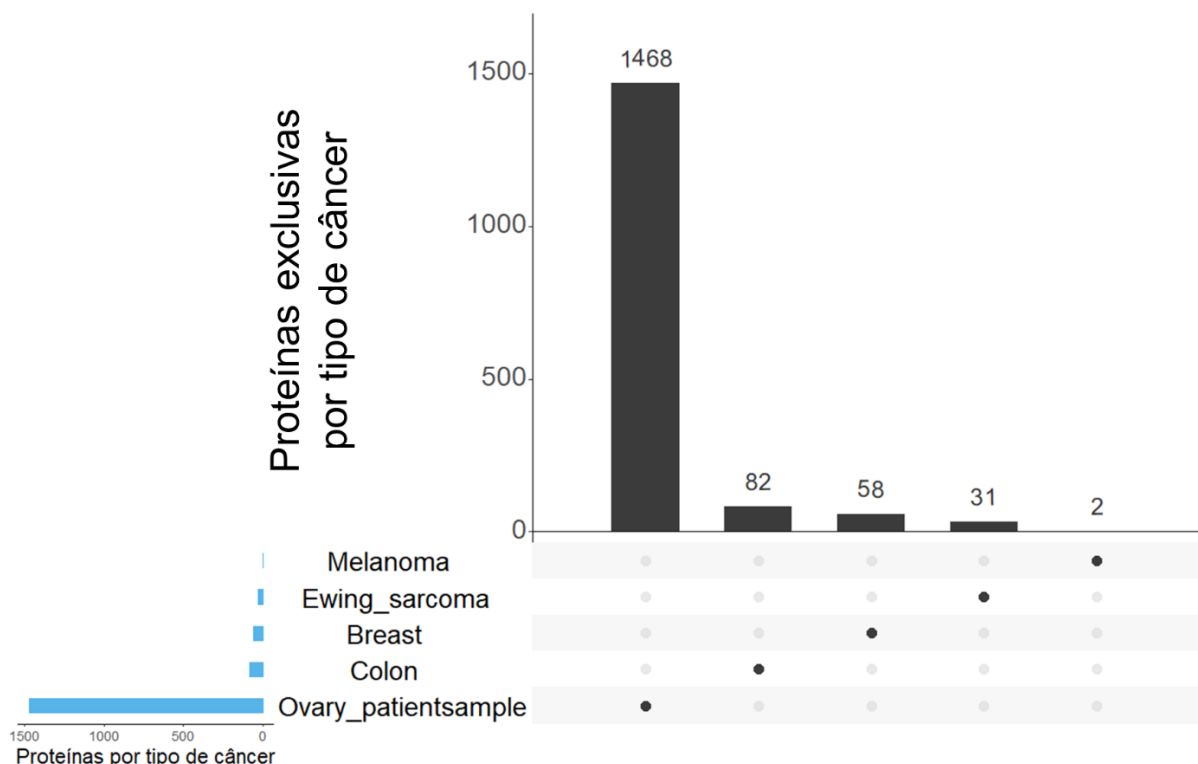


Fonte: O autor.

Nota. Panorama geral de proteínas compartilhadas em ao menos dois tipos de secretomas.

Dentre estas intersecções proteicas, observa-se 247 proteínas compartilhadas somente entre câncer de cólon e ovário e há aproximadamente 49% das 551 proteínas identificadas no secretoma tumoral de cólon também presentes no secretoma de ovário – este resultado mostra um repertório proteico comum presente em ambos os secretomas.

Também foram encontradas proteínas exclusivas para cada secretoma tumoral, característica que pode sugerir a presença de marcadores únicos (Figura 20).

Figura 20 – Proteínas exclusivas por tipo de câncer

Fonte: O autor.

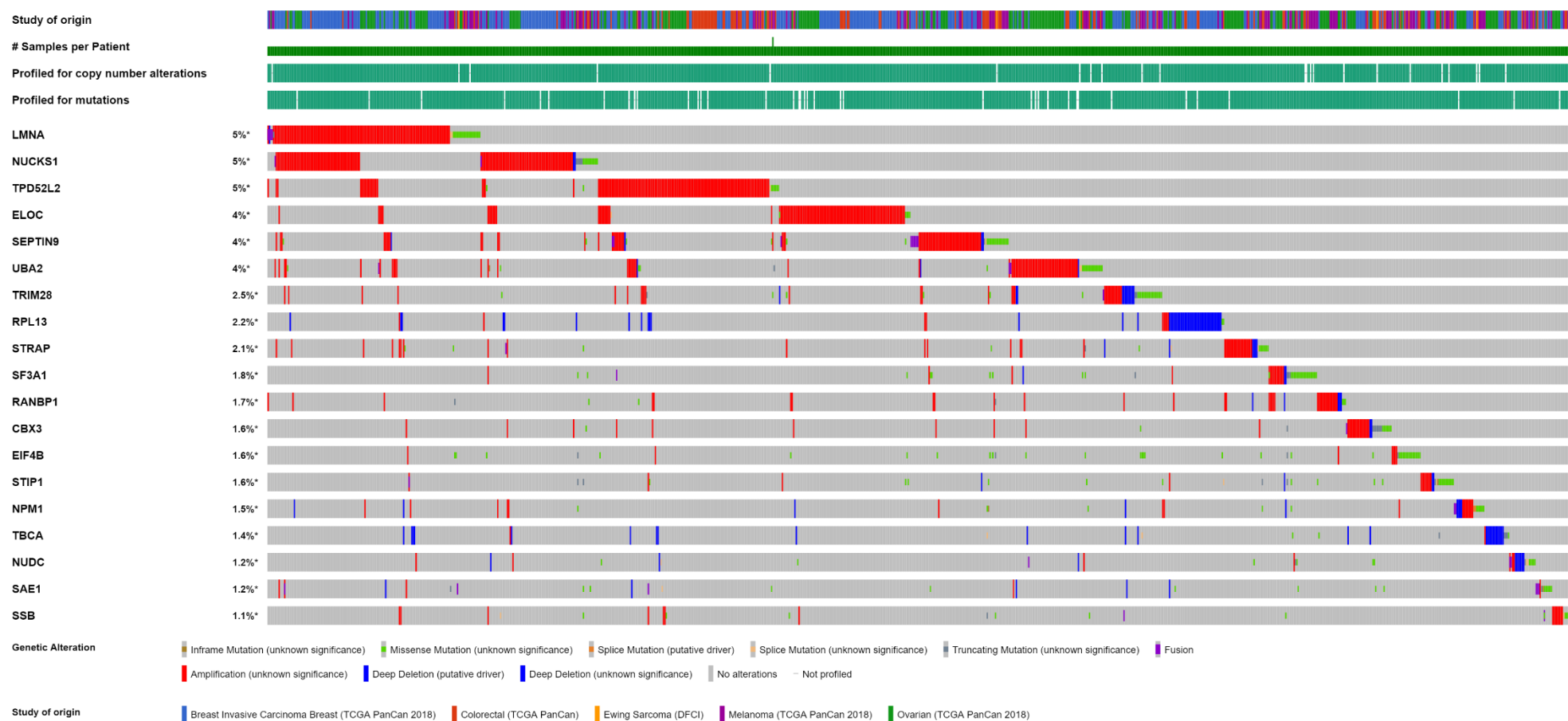
Nota. Proteínas exclusivas por tipo de câncer. O secretoma de câncer de ovário (amostra da paciente), na primeira coluna, destaca-se por fornecer um número de proteínas exclusivas maior em relação aos demais secretomas tumorais. Esta alteração deve-se provavelmente ao fato de maior disponibilidade de arquivos (RAW files) analisados em relação aos demais secretomas.

5.2.2 Alterações genéticas e perfil mutacional

Uma análise dos genes codificadores das dezenove proteínas compartilhadas foi realizada nas amostras de 2.808 pacientes com câncer de cólon, mama, ovário, melanoma e sarcoma de Ewing (Tabela 4) para verificar a ocorrência de alterações genéticas e o perfil de mutações desses genes nos diferentes tipos de câncer.

A análise relevou que 944 (34%) pacientes possuem pelo menos uma alteração em um dos dezenove genes. Observa-se também um maior número de pacientes com amplificações nos genes LMNA, NUCKS e TPD52L2 nos diferentes tipos de câncer em relação aos outros genes e alterações em 5% dos pacientes analisados (Figura 21).

Figura 21 – Perfil das alterações genéticas dos dezoito genes compartilhados em pacientes de acordo com os diferentes tipos de câncer estudados

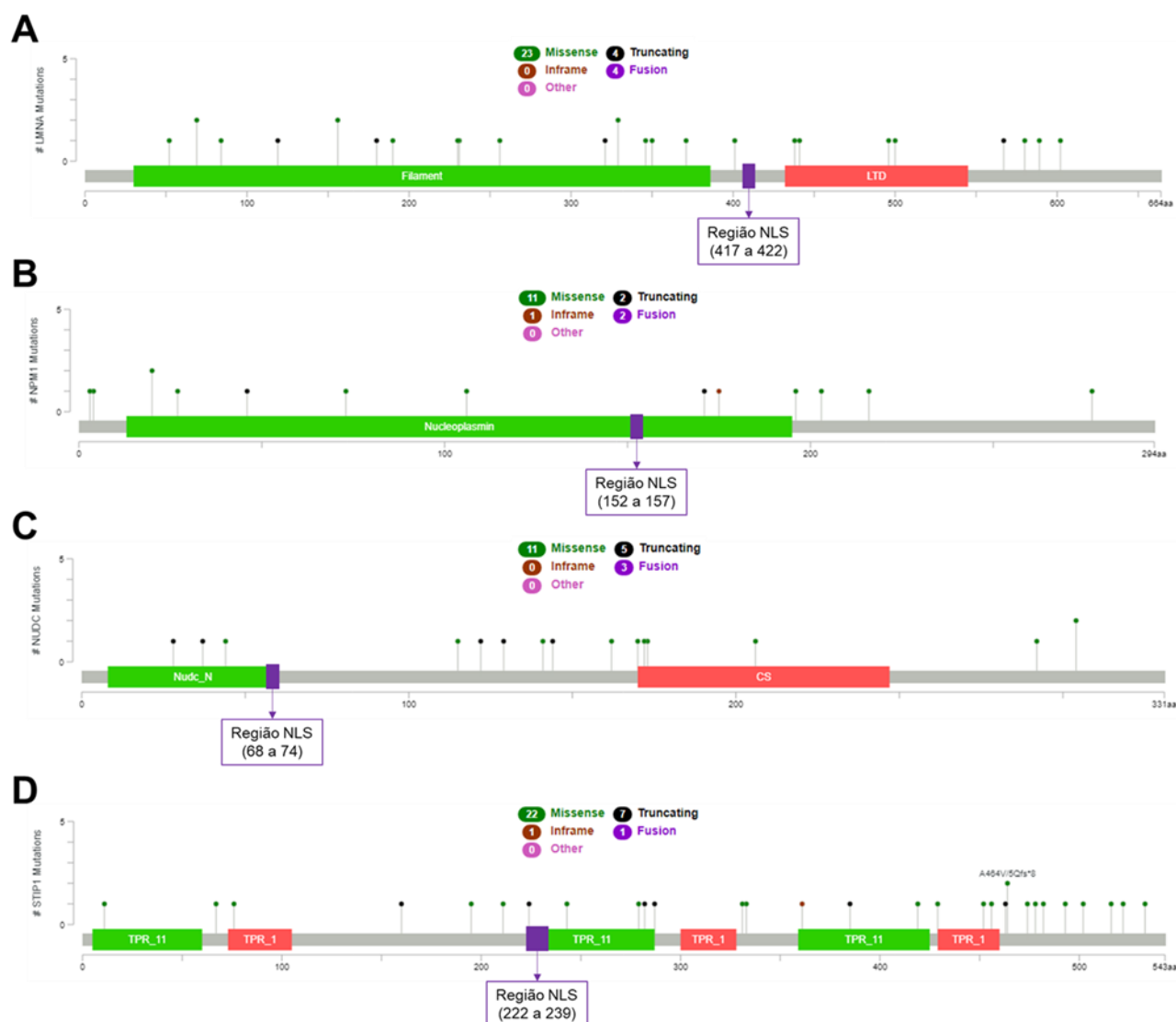


Fonte: cBioPortal.

Nota. Representação do perfil de alterações genéticas dos dezoito genes compartilhados em pacientes através de um *oncoprint*. Nesta figura estão representados somente os pacientes que possuem pelo menos uma alteração em um dos dezoito genes. Os pacientes sem qualquer alteração em pelo menos.

Destes dezenove genes, LMNA, NPM1, NUDC e STIP1 possuem o NLS identificado no banco de dados e 285 pacientes (9%) apresentaram alguma alteração em um desses quatro genes. Além disso, foi observado um número expressivo de mutações na região NLS nesses genes e para todas as mutações encontradas na região NLS, as mutações missense também se destacaram (Figura 22).

Figura 22 – Perfil mutacional dos genes LMNA, NPM1, NUDC e STIP1 em pacientes com câncer



Fonte: O Autor/cBioPortal.

Nota. Perfil de mutação dos genes codificadores das proteínas compartilhadas com NLS descrito entre os pacientes analisados. (A) Observa-se um valor expressivo de mutações no gene LMNA, no entanto, não há qualquer mutação na região NLS. (B) No gene NPM1, há um número pouco expressivo de mutações. (C) No gene NUDC também há um número pouco expressivo de mutações. (D) No gene STIP1, há alterações A464V e A465Qfs*8 em destaque, além de uma mutação X224_splice na região NLS do gene. Para todos os genes, o domínio do gene que codifica a região NLS está destacada em roxo junto com sua respectiva localização.

5.3 Localização biológica das proteínas nucleares compartilhadas em amostras histológicas

Para verificar quais proteínas compartilhadas seriam de fato nucleares em tecidos normais, foi feita uma busca de localização nas amostras histológicas normais no banco de dados *The Human Protein Atlas* em conjunto com informações relacionadas ao NLS descrito de cada proteína (se presente) no banco de dados *Uniprot/Swissprot*.

Logo, nosso critério para estabelecer uma proteína como nuclear neste trabalho foi a presença desta proteína no núcleo em amostras histológicas normais ou NLS descrito nos bancos de dados. Das dezenove proteínas compartilhadas em todos os secretomas estudados, dez foram identificadas como nucleares. A Tabela 7 resume as principais características destas proteínas.

Tabela 7 – Proteínas nucleares encontradas compartilhadas nos secretomas estudados e suas respectivas características em organismo saudável.

(continua)

Gene	Proteína	Localização em tecidos normais	NLS descrito no banco de dados	Funções primárias ¹	Referências
CBX3	<i>Chromobox protein homolog 3</i>	Núcleo.	Não.	Pertence à família da proteína Heterochomatina 1 (HP1), que está envolvida na ativação ou repressão transcricional, diferenciação, crescimento celular e modificações epigenéticas.	(MA <i>et al.</i> , 2019)
LMNA	<i>Prelamin-A/C</i>	Núcleo.	Sim – 417 a 422.	Formam uma rede de proteínas de filamento intermediário que se encontram dentro do envelope nuclear e interagem com a cromatina e o citoesqueleto; além disso, também podem fornecer uma estrutura para o envelope nuclear.	(BUTIN-ISRAELI <i>et al.</i> , 2012; IRIANTO <i>et al.</i> , 2016)
NPM1	<i>Nucleophosmin</i>	Núcleo.	Sim – 152 a 157 e 191 a 197.	Proteína nucleolar ubiquamente expressa envolvida na biogênese de ribossomos, manutenção da integridade genômica, dentre outras funções.	(ZARKA <i>et al.</i> , 2020)
NUCKS1	<i>Nuclear ubiquitous casein and cyclin-dependent kinase substrate 1</i>	Núcleo.	Não.	Desempenha um papel significativo na modulação da estrutura da cromatina e regula eventos celulares, como replicação,	(GRUNDT <i>et al.</i> , 2004; OSTVOLD;

¹ As funções primárias foram baseadas em revisões da literatura.

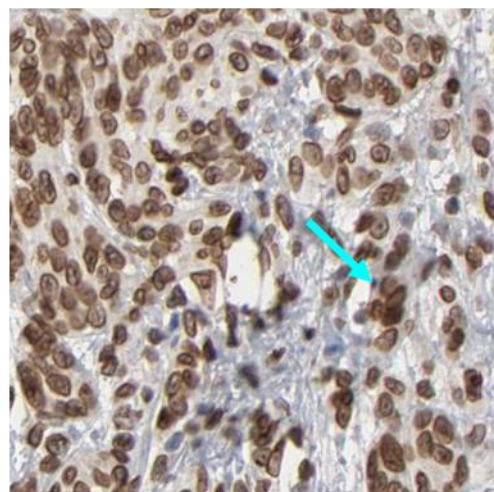
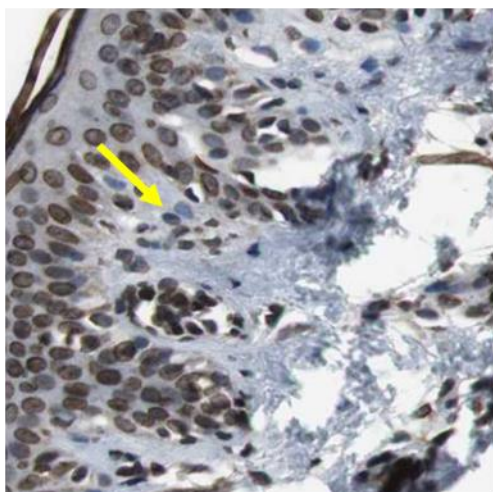
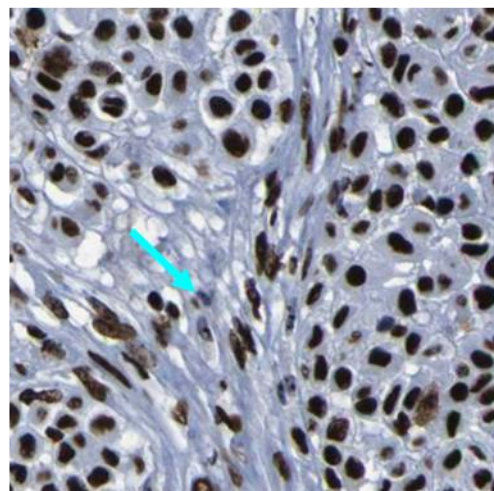
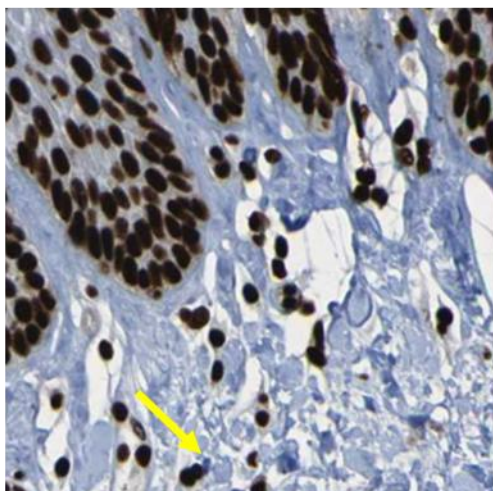

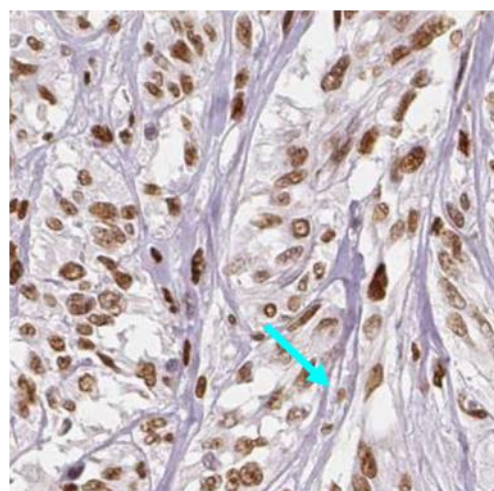
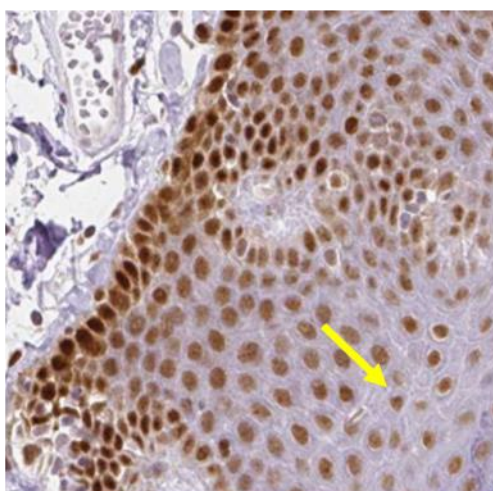
NUDC	<i>Nuclear migration protein nudC</i>	Núcleo e citoplasma.	Sim – 68 a 74.	transcrição e condensação da cromatina. Necessária para a formação correta de fusos mitóticos e separação de cromossomos durante a mitose e para citocinese e proliferação celular.	HOLTLUND; LALAND, 1985) (AUMAIS <i>et al.</i> , 2003)
SF3A1	<i>Splicing factor 3A subunit 1</i>	Núcleo.	Não.	Envolvida no processo de maturação do pré-mRNA.	(ZHAN <i>et al.</i> , 2018; ZHANG <i>et al.</i> , 2018)
SSB	<i>Ribosome-associated molecular chaperone SSB1</i>	Núcleo.	Não.	Desempenham papéis diretos na replicação, recombinação e reparo do DNA. Em muitos casos, estas proteínas formam complexos específicos com diversas proteínas de manutenção do genoma.	(MARCEAU, 2012)
STIP1	<i>Stress-induced-phosphoprotein 1</i>	Núcleo e citoplasma.	Sim – 222 a 239.	Também chamada de Hop, é uma co-chaperona que liga reversivelmente as chaperonas Hsp70 e Hsp90.	(ODUNUGA; LONGSHAW; BLATCH, 2004)
TRIM28	<i>Transcription intermediary factor 1-beta</i>	Núcleo.	Não.	Participa em muitos aspectos, tanto promovendo a proliferação celular quanto mediando atividades anti-proliferativas. Também participa da regulação da expressão gênica através da formação de heterocromatina.	(CZERWIŃSKA; MAZUREK; WIZNEROWICZ, 2017)
UBA2	<i>SUMO-activating enzyme subunit 2</i>	Núcleo.	Não.	Participa da progressão do ciclo celular, regulação da apoptose, manutenção da integridade do genoma, reparo do DNA, sobrevivência celular.	(HAY, 2005)

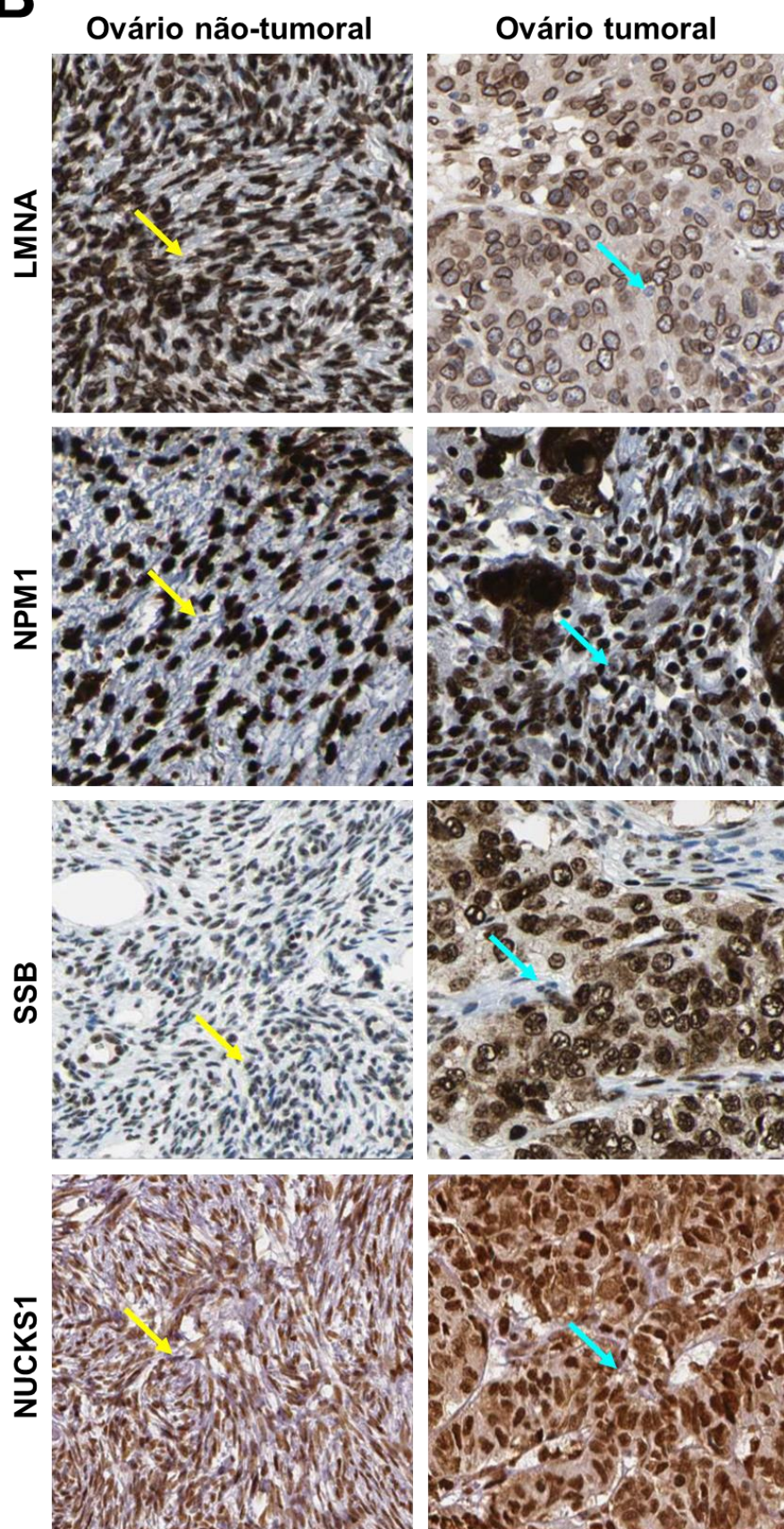
Para verificar se os dados encontrados nas análises bioinformáticas corroboram as evidências experimentais, buscas com estas dez proteínas compartilhadas nucleares foram realizadas na plataforma *The Protein Atlas*, avaliando suas localizações em tecidos não tumorais e tumorais de pacientes.

Foram analisados tecidos não tumorais de pele, cólon, ovário e mama e seus respectivos tecidos tumorais. Das dez proteínas analisadas, LMNA, NPM1, NUCKS1 e SSB mostraram alteração em sua localização celular: estas proteínas estavam presentes no núcleo em tecidos não tumorais e foram encontradas no citoplasma em tecidos tumorais (Figura 23).

As demais proteínas nucleares não mostraram alteração de localização em nenhuma amostra histológica tumoral, mantendo-se no núcleo. A análise completa da localização das dezenove proteínas compartilhadas está disponível no Anexo L deste trabalho.

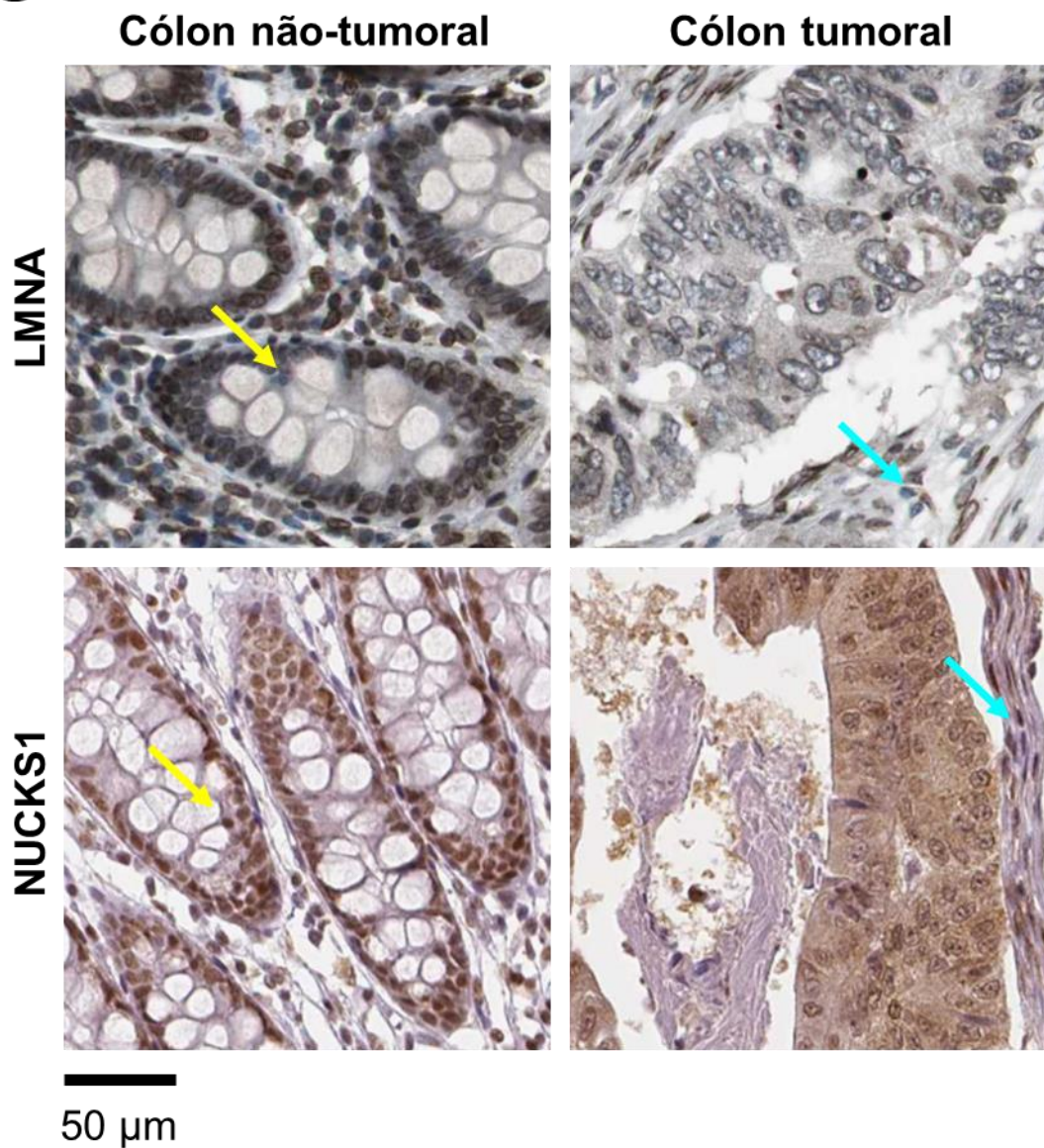
Figura 23 – Localização das proteínas compartilhadas em tecidos não tumorais e tumorais

A**Pele não-tumoral****Melanoma****LMNA****NPM1****NUCKS1**
50 μ m

B

50 μ m

C



Fonte: *The Protein Atlas*.

Nota. Localização proteínas compartilhadas com mudança de localização em tecidos não-tumorais e tumorais obtidos no banco de dados *The Protein Atlas*. Em (A-C), observa-se a presença das proteínas correspondentes no núcleo nas amostras não tumorais (seta amarela) e no citoplasma (seta azul) nas amostras tumorais. Na figura é mostrada somente as amostras histológicas onde ocorreu mudança de localização do núcleo para o citoplasma. A análise completa da localização de todas as proteínas compartilhadas está disponível no Anexo H.

6 DISCUSSÃO

Os secretomas tumorais têm sido associados continuamente com processos de desenvolvimento e progressão do câncer, contribuindo para a manutenção de um microambiente favorável ao crescimento tumoral e evidenciando que o tumor não age isoladamente durante sua progressão em tecidos e órgãos (PALTRIDGE; BELLE; KHEW-GOODALL, 2013). Deste modo, estudos do secretoma tumoral têm sido frequentemente conduzidos na busca de diferentes padrões e possíveis biomarcadores tumorais (GRØNBORG *et al.*, 2006; LAWLOR *et al.*, 2009; MATHIAS *et al.*, 2009).

Uma das observações mais intrigantes na análise de amostras de secretomas tumorais é a eventual presença de proteínas intracelulares. A presença de proteínas intracelulares é geralmente associada com contaminações derivadas das amostras ou então advindas do soro fetal bovino (nos casos de secretomas derivados de linhagens celulares em cultura); contudo, diferentes exemplos de proteínas intracelulares com funções alternativas no meio extracelular têm sido reportados (BUTLER; OVERALL, 2009; NICKEL; RABOUILLE, 2009; PUTZ *et al.*, 2012; RADISKY *et al.*, 2009). Esta característica de algumas proteínas intracelulares pode indicar uma nova explicação para sua presença no secretoma de linhagens tumorais.

6.1 Análise do repertório geral e excedente proteicos nos secretomas estudados

Nos cinco secretomas que analisamos, antes dos filtros, foram identificadas 6.092 proteínas únicas. Em secretomas normais, há diferentes predições sobre o número de proteínas secretadas, contudo, atribuindo como referência a metodologia utilizada e os dados considerados pela plataforma *The Human Protein Atlas* (UHLÉN *et al.*, 2015), estima-se que o secretoma de um organismo saudável possua aproximadamente 4.361 proteínas em sua totalidade. É importante salientar que estes dados consideram o organismo como um todo e não apenas um órgão ou um grupo de células isoladas. Portanto, ao compararmos os dados das proteínas únicas, este número mostra que o secretomas tumorais possuem mais proteínas do que o secretoma de organismos saudáveis.

Quando avaliados nossos secretomas separadamente, o número de proteínas antes dos filtros variou de 411 (melanoma) a 5.396 proteínas (ovário). Em secretomas normais, a metodologia da plataforma *The Human Protein* (UHLÉN *et al.*, 2015, 2019)

estimou um número diferente de proteínas para cada compartimento, destas, 484 foram anotadas como segregadas para compartimentos locais como tecidos masculinos ou femininos e pele, por exemplo. A principal conclusão que chegamos é que, principalmente por se tratar de uma análise local, todos os secretomas tumorais possuem um número maior de proteínas secretadas em relação aos secretomas normais ou saudáveis.

Ao avaliar o reportório proteico após o uso dos filtros, observa-se que há um excedente proteico que varia de 17% (melanoma) a 39% (ovário) quando comparado com a quantidade inicial de proteínas – antes da aplicação dos filtros. Avaliando-se somente as proteínas únicas, este excedente proteico representa 38% do secretoma (2.312 proteínas) e mostra um repertório proteico que participa de vias ligadas principalmente ao RNA, condição não prevista nas análises do secretoma normal (UHLÉN *et al.*, 2015, 2019). Portanto, este excedente proteico pode indicar que além das proteínas já eventualmente secretadas por um organismo saudável, pode haver também um reportório específico de proteínas sendo secretadas exclusivamente por células tumorais.

O repertório do secretoma tumoral de cólon mostrou aproximadamente 49% de similaridade proteica em relação ao de ovário, de modo que 247 proteínas foram identificadas em ambos os secretomas.

Alguns sobreviventes de câncer podem desenvolver um novo tipo da doença não relacionado ao câncer anterior – esta condição é chamada de segundo câncer (MORTON *et al.*, 2014). Certos tipos de câncer e tratamentos contra a doença podem estar associados a um risco maior de desenvolvimento de certos tumores secundários. Sobreviventes de câncer de ovário, por exemplo, podem ter qualquer tipo de segundo câncer, no entanto, há um risco aumentado de câncer de cólon (BOUCHARDY *et al.*, 2008; EVANS *et al.*, 2002).

Ademais, o carcinoma metastático mais comum que simula o câncer ovariano primário tem origem no intestino grosso havendo, inclusive, casos reportados em que o carcinoma ovariano metastático de origem no intestino grosso simula um câncer de ovário primário (DAYA; NAZERALI; FRANK, 1992). O contrário também fora recentemente relatado em um estudo de caso clínico, no qual uma paciente, teoricamente curada de câncer de ovário e que estava há três anos livre da doença após cirurgia e tratamento quimioterápico, apresentou um câncer de cólon metastático

de origem ovariana simulando um câncer de cólon primário (PARK; JUNG; BAEK, 2018). Estes estudos mostram que há, de fato, uma relação intrínseca entre estes dois tipos de câncer, o que pode explicar a similaridade proteica identificada dentre estes secretomas tumorais em nosso trabalho.

A análise das proteínas compartilhadas também mostrou que, dentre o excedente proteico de todos os secretomas tumorais, há dezenove proteínas presentes em todos os secretomas analisados, mostrando um repertório proteico comum em diferentes tipos de câncer.

6.2 O papel da SUMOilação em diferentes proteínas

Durante a análise destas dezenove proteínas compartilhadas, observamos principalmente a presença de proteínas oriundas do núcleo através da análise do componente celular destas proteínas. Além disso, destacou-se a presença de proteínas que participam tanto da via de SUMOilação quanto proteínas SUMOiladas.

A SUMOilação visa principalmente proteínas nucleares e afeta a regulação da transcrição, reparo de DNA e estrutura da cromatina (HOEGE *et al.*, 2002), além de regular o tráfego núcleo-citoplasmático (MATUNIS; COUTAVAS; BLOBEL, 1996).

Curiosamente, Santiago *et al.* (2013) mostraram que a localização de p53 pode estar relacionada com sua SUMOilação. No estudo experimental, eles observaram que a proteína p53 quando fundida a uma porção SUMO apresentou uma localização predominantemente citoplasmática quando compara às proteínas não fundidas à SUMO, que foram identificadas exclusivamente no núcleo. Coletivamente, os resultados do experimento apoiam a noção de que a SUMOilação de p53 facilita sua exportação nuclear, embora não tenha sido explicitado se a mudança de localização ocorreu também a partir de seu sinal de exportação (SANTIAGO *et al.*, 2013).

Na regulação do tráfego núcleo-citoplasmático, a SUMOilação também desempenha um papel crítico no sistema nuclear e localização subnuclear das proteínas. A proteína 1 ativadora da Ran GTPase (RanGAP1; do inglês *Ran GTPase-activating protein 1*) – que ativa a Ran-GTPase – foi o primeiro substrato descoberto a ser modificado por proteínas SUMO. RanGAP1 está envolvida em processos de direcionalidade de transporte nuclear e citoplasmático: RanGAP1 não modificada está localizada no citoplasma, no entanto, após a modificação por proteínas SUMO, RanGAP1 interage com a proteína de ligação específica ao Ran 1 (RanBP1; do inglês,

Ran binding protein 1) ou a proteína de ligação específica ao Ran 2 (RanBP2; do inglês, Ran binding protein 2), proteínas do complexo de poro nuclear (NPC) (PICHLER *et al.*, 2002). RanGAP1 sozinha não pode ativar a hidrólise de GTP, portanto coopera com RanBP1 ou RanBP2 para ativar a GTPase nos complexos contendo RanGTP (MATUNIS; COUTAVAS; BLOBEL, 1996).

RanBP1 foi identificada em todos os secretomas mas em condições normais funciona como uma coativadora de RanGAP1, sendo localizada no citosol. Não sabemos como esta proteína pode ser secretada, contudo, é possível que ela interfira tanto no mecanismo de exportação nuclear, sendo também responsável para exportação de proteínas nucleares no citosol, quanto na secreção das proteínas nucleares.

A SUMOilação anormal pode levar ao desenvolvimento de uma série de doenças, incluindo câncer. Muito embora a associação entre a expressão de vários componentes na via de sinalização SUMO e a progressão tumoral ou metástase ainda não seja totalmente compreendida, um número crescente de estudos tem mostrado que a SUMOilação desempenha um papel importante no câncer (BETTERMANN *et al.*, 2012; EIFLER; VERTEGAAL, 2015; FLOTHO; MELCHIOR, 2013; RABELLINO; ANDREANI; SCAGLIONI, 2017; SEELER; DEJEAN, 2017). Uma das hipóteses, por exemplos, é que a SUMOilação possa contribuir substancialmente para a sobrevivência e proliferação das células cancerosas em um microambiente potencialmente hostil (SEELER; DEJEAN, 2017).

Em um recente estudo, Neo *et al.* (2015) mostraram que TRIM28 é uma ligase E3 SUMO responsável pela SUMOilação, mediada por ARF, de NPM1. ARF e TRIM28 cooperam para SUMOilar NPM1, proteína nucleolar que regula a duplicação do centrôssomo e estabilidade genômica. Neste estudo, a SUMOilação mediada por ARF de NPM1 foi atenuada com a falta de TRIM28 e aprimorada pela superexpressão de TRIM28. Com isso, a co-expressão de ARF e TRIM28 promoveu a localização centrôssomal NPM1, aumentando sua SUMOilação e suprimindo a amplificação do centrôssomo; essas funções exigiam a atividade da ligase E3 de TRIM28 (NEO *et al.*, 2015).

Por outro lado, a depleção de ARF ou TRIM28 aumentou a amplificação do centrôssomo. A amplificação do centrôssomo é frequentemente induzida por processos oncogênicos, levando à instabilidade genômica, portanto, o aumento da

SUMOilação de NPM1 por ARF/TRIM28 contribuiu para sua localização centrossômica aprimorada e supressão da amplificação do centrossomo, sugerindo uma nova função supressora de tumor independente de p53 (NEO *et al.*, 2015).

A identificação tanto de NPM1 quanto TRIM28 em todos os nossos secretomas tumorais pode indicar que NPM1 não fora SUMOilada por TRIM28, levando à instabilidade cromossômica em todos os tipos de câncer.

Além das proteínas compartilhadas diretamente relacionadas a mecanismos de transporte nuclear e SUMOilação, também foram identificados em todos os secretomas demais proteínas nucleares. Em nossos resultados, as dezenove proteínas compartilhadas identificadas também mostraram possuir pelo menos um sítio de SUMOilação, o que indica que todas estas proteínas podem ser SUMOiladas. Ademais, estas proteínas podem não só participar de mecanismos distintos de secreção não convencional, mas também desempenhar diferentes funções, pois a identificação de funções secundárias em proteínas intracelulares quando secretadas – principalmente em processos patológicos como o câncer – não é incomum (MIN; LEE; BAEK, 2016).

6.3 Evidências de proteínas intracelulares secretadas e seu novo papel em patogêneses

Uma das proteínas com múltiplas funções encontradas nos secretomas analisados foi a *stress-induced phosphoprotein 1* (STIP1); esta proteína foi identificada em nosso estudo em todos os secretomas tumorais.

STIP1 é uma co-chaperona que liga reversivelmente as proteínas chaperonas Hsp70 e Hsp90 (ODUNUGA; LONGSHAW; BLATCH, 2004). Ela se localiza tanto no núcleo quanto no citoplasma e sua sequência NLS permite que ela se desloque do núcleo para o citoplasma através do controle de quinases (LONGSHAW *et al.*, 2004). Contudo, mesmo não possuindo qualquer peptídeo sinal, ela também pode ser translocada para a superfície celular ou secretada para fora da célula (EUSTACE; JAY, 2004; WANG *et al.*, 2010).

No estudo *in vitro* de TSAI *et al.* (2012), foi observado que a proteína STIP1, quando secretada por células de câncer de ovário, fora capaz de estimular a proliferação celular através da ativação da expressão gênica de ID3 (*inhibitor of DNA binding 3*) (TSAI *et al.*, 2012).

Em outro estudo, MA *et al.* (2020) mostraram que STIP1 pode ser secretada por células do carcinoma hepatocelular (um tipo de câncer primário do fígado), podendo agir como um marcador da progressão da doença. Neste estudo, foram analisadas amostras séricas de 340 pacientes com carcinoma hepatocelular que passaram por diferentes tratamentos de quimioembolização e ressecção do tumor. As amostras dos pacientes foram analisadas para avaliar o papel de STIP1 como biomarcador no prognóstico da doença e observou-se que pacientes com bom prognóstico apresentaram baixos níveis séricos de STIP1, enquanto pacientes com um pior prognóstico da doença apresentaram altos níveis séricos de STIP1, mesmo após tratamento (MA *et al.*, 2020). Desta forma, demonstrou-se que a proteína STIP1 pode ser um biomarcador promissor da progressão desse tipo de câncer. Considerando os achados para carcinoma hepatocelular e os resultados encontrados, sugere-se que STIP1 possa atuar como um marcador de prognóstico para os tipos de câncer avaliados.

Além de estudos relativos ao câncer, há diversos estudos relacionando as consequências da proteína STIP1 secretada na interação com a proteína príon celular (PrP^C) em doenças neurodegenerativas (ARANTES *et al.*, 2009; BERALDO *et al.*, 2010, 2013; CAETANO *et al.*, 2008; HAJJ *et al.*, 2013; LOPES *et al.*, 2005; MACIEJEWSKI *et al.*, 2016; MARTINS *et al.*, 2002; OSTAPCHENKO *et al.*, 2013; ROFFÉ *et al.*, 2010).

A proteína STIP1 deve ser secretada para interagir com PrP^C na superfície celular e assim como Hsp70 e Hsp90, STIP1 também pode ser secretada através de exossomos (CLAYTON *et al.*, 2005; HAJJ *et al.*, 2013; LANCASTER; FEBBRAIO, 2005), podendo então se ligar à superfície de neurônios, interferindo, juntamente com PrP^C, em uma série de eventos ligados ao crescimento e sobrevivência celular. Deste modo, seus mecanismos de sinalização e secreção em células cancerígenas podem também ser similares àsquelas observadas nos neurônios.

Outra proteína nuclear que também foi identificada nos secretomas analisados foi a proteína *high mobility group A1* (HMGA1); esta proteína foi encontrada em quatro dos cinco secretomas estudados, não sendo encontrada apenas no secretomas do melanoma.

A proteína HMGA1 é expressa em altos níveis durante a embriogênese mas sua expressão é baixa ou ausente em tecidos adultos diferenciados (CHIAPPETTA *et*

al., 1996). No entanto, sua expressão volta a ser elevada em adultos quando associada com o início de diferentes tipos de câncer, incluindo câncer de pulmão (HILLION *et al.*, 2009; KETTUNEN *et al.*, 2004), cólon (WILLIAMS *et al.*, 2015), glioblastoma (COLAMAIIO *et al.*, 2016), osteosarcoma (XU *et al.*, 2014) e câncer de mama (HUANG *et al.*, 2015).

No câncer de mama, o nível de expressão da proteína HMGA1 indica um resultado adverso no prognóstico clínico. Ademais, no câncer de mama com subtipo triplo negativo também há altos níveis de expressão de HMGA1 (SHUCH, BRIAN; LINEHAN, B. W. M.L.; SRIVASAN, 2012).

Ao estudarem as funções de HMGA1 no câncer de mama triplo-negativo, MENDEZ *et al.* (2018) mostraram que dependendo de sua localização celular, núcleo ou citoplasma, esta proteína pode induzir uma maior malignidade e metástase tumoral. A expressão de HMGA1 por células invasivas de câncer de mama indica a possibilidade da proteína estabelecer novas interações moleculares no espaço extracelular, pois sua secreção se correlaciona com um fenótipo invasivo em pacientes com câncer de mama triplo-negativo (MENDEZ *et al.*, 2018).

Tanto a proteína STIP1 quanto HMGA1 não possuem peptídeo sinal, portanto, não entram no mecanismo convencional de secreção proteica que se utiliza da rota RE–Golgi. MENDEZ *et al.* (2018) mostraram que a variante secretada de HMGA1, em uma linhagem mais invasiva derivada de MDAMB231, não é secretada pela via tradicional RE–Golgi. Contudo, as vias de secreção não convencionais tanto para STIP1 quanto HMGA1 ainda não foram completamente elucidadas em diferentes tipos de câncer e permanecem incertas (KLUNE *et al.*, 2008; MÉNDEZ *et al.*, 2019).

Outra proteína encontrada no secretoma que apresentou mudança de localização do núcleo para o citoplasma nos tecidos analisados é a *nucleophosmin 1* (NPM1). Esta proteína também apresentou uma importante mudança de localização quando comparados tecidos não tumorais de pele e ovário com tecidos tumorais de pacientes. Em tecidos não tumorais ela foi identificada no núcleo; entretanto, nos tecidos de melanoma e câncer de ovário, ela teve sua localização alterada para o citoplasma.

A NPM1 é uma proteína nucleolar ubiquamente expressa e apesar de residir no nucléolo em condições fisiológicas normais, continuamente se move do núcleo para o citoplasma para desempenhar suas inúmeras funções como chaperona

(ZARKA *et al.*, 2020). Esta proteína também é considerada uma proteína *hub* no nucléolo, onde ajuda a manter a organização nucleolar, pois é altamente conectada a outras proteínas (DI MATTEO *et al.*, 2017).

O domínio C terminal desta proteína possui uma região altamente conservada, onde se encontra seu sinal de localização nucleolar (NoLS; do inglês, *nucleolar localization signal*) (FEDERICI; FALINI, 2013; NISHIMURA *et al.*, 2002). A região NoLS é crítica para a localização de NPM1 no nucleoplasma e qualquer alteração neste local é suficiente para mudar a localização de NPM1 para fora do nucléolo (NISHIMURA *et al.*, 2002).

Mutações no NoLS causam uma mudança de localização da proteína NPM1 e quando ocorrem ela não vai para o núcleo e passa a permanecer exclusivamente no citoplasma. Esta característica é única para a leucemia mieloide aguda (LMA), doença caracterizada por uma expansão clonal e alta proliferação de células mieloides na medula óssea, sendo encontrada em 20–30% de todos os novos casos diagnosticados com a doença (PAPAEMMANUIL *et al.*, 2016).

A proteína NPM1 citoplasmática (NPM1c) é detectada na LMA, assim como seu gene NPM1 mutado, e não existem mutações no gene NPM1 capazes de fazer com que a proteína NPM1 ainda permaneça no nucléolo. Curiosamente, a proteína NPM1c também é capaz de recrutar a proteína NPM1 sem alterações para o citoplasma e perturbar sua função normal (GRISENDI *et al.*, 2005).

Em nosso estudo não foram encontradas mutações exclusivas nas regiões na proteína NPM1 nos tipos de câncer estudados; contudo, mais estudos são necessários para identificar a influência de mutações na proteína em outros tipos e câncer.

LMNA codifica a proteína lamina A/C. Esta proteína também se destacou nas análises, sendo encontrada nos secretomas de todos os tipos de câncer. Além disso, teve sua mudança de localização confirmada do núcleo para o citoplasma nas amostras histológicas de câncer de cólon, câncer de ovário e melanoma que analisamos.

Laminas formam uma rede de proteínas de filamento intermediário que se encontram dentro do envelope nuclear e interagem com a cromatina e o citoesqueleto; acredita-se também que elas podem fornecer uma estrutura para o envelope nuclear (IRIANTO *et al.*, 2016). Em seres humanos, as principais formas da proteína lamina

são expressas a partir de três genes: as laminas A e C são produtos alternativos de splicing do gene LMNA (coletivamente chamadas laminas tipo A) e as laminas B1 e B2 são codificadas pelos genes LMNB1 e LMNB2 (laminas tipo B) (IRIANTO et al., 2016).

Recentemente, laminas têm sido estudadas como potenciais mecanossensores que recebem informações e respondem a mudanças dependendo das propriedades estruturais e mecânicas de substratos celulares e do microambiente ao qual são inseridas (ISERMANN; LAMMERDING, 2013; SWIFT et al., 2013). Um estudo recente mostrou que células cultivadas em substratos mais macios têm núcleos mais macios enquanto células que crescem em substratos mais rígidos têm núcleos mais rígidos, de modo que as mudanças na rigidez nuclear foram relacionadas aos níveis da lamina A (SWIFT et al., 2013).

Um processo fisiológico relevante que é dependente da rigidez nuclear é a metástase tumoral. Durante a migração de células tumorais em tecidos sólidos, o núcleo é submetido a forte estresse mecânico. As células com altos níveis de lamina A têm núcleos rígidos que não podem ser deformados facilmente, dificultando a migração. Em contrapartida, as células com baixos níveis de lamina A têm núcleos flexíveis e podem se espremer facilmente através do tecido durante a migração. Os núcleos flexíveis, no entanto, são propensos a fragmentação levando à morte celular, de modo que a quantidade certa de lamina A no núcleo seja crucial para garantir uma migração bem sucedida e a sobrevivência das células (HARADA et al., 2014).

Núcleos mais maleáveis são necessários para a migração das células tumorais pois assim não se romperiam em consequência do estresse mecânico sofrido pela célula. As laminas são proteínas que formam a estrutura do núcleo. De fato, para a célula migrar e o núcleo não se romper, ele necessariamente precisaria ter uma quantidade menor de laminas em relação às células normais. Sugerimos, portanto, que se as laminas não forem encaminhadas para o núcleo, necessariamente este será menos rígido, pois terá uma menor quantidade de proteínas estruturais que o tornariam menos flexível.

Embora pouco se saiba sobre as laminas durante a migração celular e seu papel no meio extracelular, sugerimos esta proteína não seja enviada para o núcleo para promover uma maior flexibilidade nuclear, facilitando a migração celular. No entanto, mais estudos são necessários para elaborar a ligação entre as laminas e o

câncer, o que pode levar a novos tratamentos ou ao menos uma compreensão mais clara das laminas na progressão tumoral.

A proteína *chromobox protein homolog 3* (CBX3) é um gene prognóstico desfavoráveis no câncer de mama e foi identificada dentre as dezenove proteínas compartilhadas em nosso trabalho. Esta proteína é codificada pelo gene CBX3 e é uma componente da heterocromatina, portanto é importante no silenciamento transcricional (RUSSO *et al.*, 2007). Ademais, a coloração imuno-histoquímica usando um anticorpo direcionado a CBX3 mostra um padrão de expressão diferencial em amostras de câncer de mama (UHLÉN *et al.*, 2015). Em nossas análises, outras proteínas também mostraram diferentes padrões de expressão quando comparados os tecidos normais e tumorais, o que indica que elas podem ser mais ou menos expressas nos diferentes tipos de câncer quando comparadas com tecidos normais.

Uma análise do perfil de alterações genéticas também foi realizada e a presença de amplificações se destacou ao serem analisados os genes codificadores das dezenove proteínas compartilhadas em amostras de pacientes na plataforma *cBioPortal*. Contudo, ainda não há evidências suficientes que associem as mutações, principalmente *missense*, com a deslocalização das proteínas e eventuais mudanças de funções. Este estudo permanece, portanto, a ser explorado com maior rigor no futuro.

Além de investigar sobre como as proteínas podem ter sua localização alterada do núcleo para o citosol e do citosol para o meio extracelular, também é essencial discutir sobre quais mecanismos podem ser responsáveis por desviar essas proteínas de vias de secreção convencionais e direcioná-las para meios alternativos de liberação no meio extracelular, a partir de exemplos de mecanismos não convencionais de secreção recém elucidados.

A maioria das vias de secreção não convencionais é estimulada pelo estresse, contudo, a proteína FGF2, embora seja secretada por uma via de secreção não convencional, não é secretada fora da rota RE–Golgi em consequência do estresse celular (LA VENUTA *et al.*, 2015; STERINGER; MÜLLER; NICKEL, 2015).

Esta proteína induz a formação de um poro que a permite se translocar através da membrana plasmática para alcançar o espaço extracelular, no entanto, ainda não está claro se o poro é formado em um lado da célula e logo após do outro lado, ou se

o poro permanece estável para permitir a translocação de FGF2 (LA VENUTA *et al.*, 2015; STERINGER; NICKEL, 2018).

Quando FGF2 foi forçada a usar a via de secreção convencional, através da adição de um peptídeo sinal a ela, a proteína recebeu modificações indesejadas ao passar pelo controle de qualidade de Golgi e foi secretada em uma forma biologicamente não ativa (WEGEHINGEL; ZEHE; NICKEL, 2008). Além disso, como a translocação de FGF2 é estritamente dependente de seu enovelamento, a via de secreção alternativa pode constituir um mecanismo de controle de qualidade próprio, que permite que o enovelamento da proteína seja o ideal para sua translocação (TORRADO *et al.*, 2009).

A proteína FGF2 ilustra o acionamento de um mecanismo de secreção não convencional para não a direcionar aos controles convencionais de RE–Golgi, pois estes modificariam a proteína e a tornariam inativa no meio extracelular. Ao não ser direcionada para a rota RE–Golgi, a proteína é capaz de manter sua completa integridade e funcionalidade fora da célula.

6.4 Considerações finais

A coleção de proteínas secretadas por uma célula – o secretoma – é de particular interesse na fisiopatologia do câncer devido ao seu potencial diagnóstico e papel na tumorigênese. No entanto, estudos de secretoma tumoral são frequentemente limitados a um tecido ou tipo de câncer, ou então a proteínas já secretadas, não explorando a potencialidade de proteínas não convencionais quando secretadas e suas funções alternativas associadas ao contexto tumoral.

De fato, as proteínas intracelulares identificadas, principalmente NPM1, STIP1, LMNA e HMGA1 já detêm evidências de possuírem um papel intracelular ativo na tumorigênese. No entanto, a participação destas proteínas no meio extracelular ainda permanece pouco ou não explorada. Neste contexto, nossos dados sugerem que estas proteínas são secretadas por vias não convencionais e, ainda, que essas vias devem ser recorrentes em diferentes tipos de câncer.

A partir de nossos resultados, podemos identificar que o número de proteínas, que são possivelmente proteínas de via secretória não canônica, varia e, dependendo do tipo de câncer, esta alteração pode variar de 17% a 39%. Dentre a porção do secretoma que contém estas proteínas, foram identificadas similaridades no repertório

proteico, mesmo em diferentes tipos de câncer, por exemplo, cólon e ovário, o que pode indicar o acionamento de mecanismos semelhantes em diferentes tumores.

A partir da identificação das proteínas excedentes compartilhadas em todos os secretomas, propomos que modificações pós-traducionais possam ser fundamentais para a modificação de mecanismos de transporte nuclear a partir da alteração do controle de SUMOilação e deSUMOilação das proteínas, o que faz com que determinadas proteínas nucleares sejam exportadas para fora do núcleo. A presença, em todos os secretomas, de UBA2 e SAE1 pode indicar que não houve a formação da enzima ativadora de SUMO para a SUMOilação de proteínas ou que esta ocorreu de forma alterada, o que poderia vir a comprometer e até mesmo alterar a interação de proteínas SUMO no transporte nuclear, já que RanGAP1 precisa passar pelo processo de SUMOilação para dar direcionalidade no transporte. Ademais, a exemplo da proteína p53, é possível também que as proteínas nucleares sejam recrutadas e exportadas para fora do núcleo mesmo sem um sinal de exportação nuclear a partir da interação com proteínas SUMO – principalmente porque as dezenove proteínas compartilhadas possuem sítios de SUMOilação. Além disso, a presença de RanBP1 em todos os secretomas pode indicar uma participação ativa não só na direcionalidade do transporte nuclear, mas também na secreção das demais proteínas; não excluimos, portanto, que RanBP1 também possa desempenhar outras funções em um cenário tumoral. Portanto, estudos direcionados seriam necessários para a elucidação da participação de RanBP1 tanto intracelular quanto no microambiente tumoral.

Não sabemos ao certo se estas proteínas intracelulares são secretadas em consequência de sinalizações intracelulares ou extracelulares, no entanto, apesar de não possuírem um peptídeo sinal, é possível que estas proteínas apresentem modificações para que sejam recrutadas e encaminhadas para fora da célula por vias não convencionas. Sugerimos, portanto, que essa identificação possa ser derivada de uma mudança estrutural determinante na proteína a partir de modificações em sua estrutura primária ou inseridas na proteína a partir, por exemplo, da SUMOilação.

Após essa modificação na proteína intracelular para que ela seja secretada, a proteína provavelmente não seria encaminhada para a via convencional de secreção em Golgi, pois, a exemplo da proteína FGF2, ela poderia passar por modificações que a levariam para degradação ou a tornariam inativa fora da célula, de modo que sua secreção não interfira no processo da tumorigênese. Reforçamos que se trata de uma

hipótese baseada em nossos resultados a partir da análise dos dados e que seria imprescindível um estudo *in vitro* focado na relação entre mecanismos de transporte nuclear e SUMOilação para confirmar ou rejeitar esta hipótese.

Ademais, devido ao metabolismo tumoral exacerbado, também é possível que a célula cancerosa perca alguns pontos de controle e não consiga gerenciar adequadamente o envio correto de proteínas para os compartimentos celulares adequados. Neste cenário, algumas proteínas nucleares poderiam não ser encaminhadas para o núcleo da célula e permaneceriam no citosol onde, por meio de novas interações, teriam a probabilidade de seguir em rotas alternativas, inclusive para o meio extracelular. Além disso, os dois cenários poderiam ocorrer concomitantemente dentro da célula tumoral.

Por fim, é importante ressaltar que este trabalho possui limitações, pois trata-se de uma análise bioinformática do conteúdo de secretomas tumorais, havendo a necessidade de validação de nossos dados *in vitro* e *in vivo* para a elucidação de novos mecanismos de secreção proteica em células tumorais e identificação de potenciais biomarcadores séricos.

7 CONCLUSÕES

Ao longo deste trabalho foram re-analizados arquivos públicos correspondentes a corridas de LC-MS/MS, oriundos de diferentes secretomas tumorais, através de busca em bancos de dados e desenvolvimento de *scripts* para melhor visualização e análise dos dados. Nossos resultados permitiram elaborar as seguintes conclusões:

1. A análise de diferentes secretomas tumorais, obtidos e processados em diferentes laboratórios e equipamentos, apresentou uma similaridade qualitativa no repertório proteico e mostrou que a secreção proteica de células cancerígenas possui uma grande quantidade de proteínas intracelulares que possivelmente são secretadas por vias não convencionais;
2. A análise da composição dos diferentes secretomas tumorais mostrou que somente parte deles contém proteínas de possível via secretória não-canônica e, dependendo do tipo de câncer, esta porção pode variar de 17% a 39% em relação aos secretomas pré-filtragem. Para estas proteínas, foram identificadas similaridades, principalmente dentre os cânceres de cólon e ovário, o que pode indicar o acionamento de mecanismos semelhantes em diferentes tumores;
3. Dentre as proteínas secretadas, há uma quantidade considerável de proteínas oriundas do núcleo, as quais podem ter sido identificadas nos secretomas não por motivo de contaminação das amostras, mas sim por participarem de mecanismos não convencionais de secreção, possivelmente desempenhando importantes papéis no desenvolvimento e progressão tumoral;
4. A análise das funções moleculares das proteínas compartilhadas mostrou que as proteínas secretadas participam de processos de SUMOilação, uma modificação pós-traducional diretamente relacionada com a direcionalidade do transporte nuclear e que pode estar envolvida na exportação e secreção das proteínas nucleares a partir de alterações nas proteínas SUMO;

5. Por fim, é necessário ressaltar que este trabalho possui limitações, pois trata-se de uma análise bioinformática do repertório de proteínas de secretomas tumorais, havendo a necessidade de validação de nossos dados para a confirmação de nossa hipótese.

8 DISPONIBILIDADE DOS DADOS

Todos os arquivos suplementares, assim como todos os códigos escritos em R para a geração das imagens deste trabalho estão disponíveis em um repositório aberto na plataforma *GitHub* (<https://github.com/julianademorais/cancer-secretome>).

REFERÊNCIAS

- AGRAWAL, Ganesh Kumar *et al.* Plant secretome: unlocking secrets of the secreted proteins. **Proteomics**, Germany, v. 10, n. 4, p. 799–827, 2010.
- ALBERTS, B.; BRAY, D.; LEWIS, J.; RAFF, M.; ROBERTS, K.; WATSON, J. D. **Biologia Molecular da Célula**. 6. ed. Porto Alegre [s. n.].
- ALMAGRO ARMENTEROS, José Juan *et al.* SignalP 5.0 improves signal peptide predictions using deep neural networks. **Nature biotechnology**, United States, v. 37, n. 4, p. 420–423, 2019.
- ANDREI, Cristina *et al.* The secretory route of the leaderless protein interleukin 1 β involves exocytosis of endolysosome-related vesicles. **Molecular Biology of the Cell**, [S. l.], v. 10, n. 5, p. 1463–1475, 1999.
- ARANTES, Camila *et al.* Prion protein and its ligand stress inducible protein 1 regulate astrocyte development. **Glia**, [S. l.], v. 57, n. 13, p. 1439–1449, 2009.
- ARISPE, Nelson; DOH, Michael; DE MAIO, Antonio. Lipid interaction differentiates the constitutive and stress-induced heat shock proteins Hsc70 and Hsp70. **Cell stress & chaperones**, [S. l.], v. 7, n. 4, p. 330–338, 2002.
- ASHBURNER, Michael *et al.* The Gene Ontology Consortium, Michael Ashburner¹, Catherine A. Ball³, Judith A. Blake⁴, David Botstein³, Heather Butler¹, J. Michael Cherry³, Allan P. Davis⁴, Kara Dolinski³, Selina S. Dwight³, Janan T. Eppig⁴, Midori A. Harris³, David P. Hill⁴, Laurie Is. **Nature genetics**, [S. l.], v. 25, n. 1, p. 25–29, 2000.
- AUMAIS, Jonathan P. *et al.* Role for NudC, a dynein-associated nuclear movement protein, in mitosis and cytokinesis. **Journal of cell science**, England, v. 116, n. Pt 10, p. 1991–2003, 2003.
- BATEMAN, Alex. UniProt: A worldwide hub of protein knowledge. **Nucleic Acids Research**, [S. l.], v. 47, n. D1, p. D506–D515, 2019.
- BERALDO, Flavio H. *et al.* Role of $\alpha 7$ nicotinic acetylcholine receptor in calcium signaling induced by prion protein interaction with stress-inducible protein. **Journal of Biological Chemistry**, [S. l.], v. 285, n. 47, p. 36542–36550, 2010.
- BERALDO, Flavio H. *et al.* Stress-inducible phosphoprotein 1 has unique cochaperone activity during development and regulates cellular response to ischemia via the prion protein. **FASEB Journal**, [S. l.], v. 27, n. 9, p. 3594–3607, 2013.
- BERTHENET, Kevin *et al.* Extracellular HSP110 skews macrophage polarization in colorectal cancer. **Oncoimmunology**, [S. l.], v. 5, n. 7, p. e1170264, 2016.
- BETTERMANN, Kira *et al.* SUMOylation in carcinogenesis. **Cancer letters**, Ireland, v. 316, n. 2, p. 113–125, 2012.
- BHANDARI, Vinayak *et al.* Molecular landmarks of tumor hypoxia across cancer

types. **Nature Genetics**, [S. l.], v. 51, n. 2, p. 308–318, 2019. Disponível em: <http://dx.doi.org/10.1038/s41588-018-0318-2>

BOUCHARDY, Christine *et al.* Increased risk of second cancer among patients with ovarian borderline tumors. **Gynecologic oncology**, United States, v. 109, n. 2, p. 210–214, 2008.

BRAAKMAN, Ineke; BULLEID, Neil J. Protein folding and modification in the mammalian endoplasmic reticulum. **Annual review of biochemistry**, United States, v. 80, p. 71–99, 2011.

BRODSKY, Jeffrey L.; SKACH, William R. Protein folding and quality control in the endoplasmic reticulum: Recent lessons from yeast and mammalian cell systems. **Current opinion in cell biology**, [S. l.], v. 23, n. 4, p. 464–475, 2011.

BROWN, Kristy J. *et al.* Advances in the proteomic investigation of the cell secretome. **Expert review of proteomics**, [S. l.], v. 9, n. 3, p. 337–345, 2012.

BURGESS, T. L.; KELLY, R. B. Constitutive and regulated secretion of proteins. **Annual review of cell biology**, United States, v. 3, p. 243–293, 1987.

BUTIN-ISRAELI, Veronika *et al.* Nuclear lamin functions and disease. **Trends in genetics : TIG**, [S. l.], v. 28, n. 9, p. 464–471, 2012.

BUTLER, Georgina S.; OVERALL, Christopher M. Locations Complicates Drug Targeting. **Discovery**, [S. l.], v. 8, n. deCeMber, p. 935–948, 2009. Disponível em: <http://dx.doi.org/10.1038/nrd2945>

CACCIA, Dario *et al.* Secretome compartment is a valuable source of biomarkers for cancer-relevant pathways. **Journal of proteome research**, United States, v. 10, n. 9, p. 4196–4207, 2011.

CAETANO, Fabiana A. *et al.* Endocytosis of prion protein is required for ERK1/2 signaling induced by stress-inducible protein 1. **Journal of Neuroscience**, [S. l.], v. 28, n. 26, p. 6691–6702, 2008.

CARBON, S. *et al.* The Gene Ontology Resource: 20 years and still GOing strong. **Nucleic Acids Research**, [S. l.], v. 47, n. D1, p. D330–D338, 2019.

CERAMI, Ethan *et al.* The cBio Cancer Genomics Portal: An open platform for exploring multidimensional cancer genomics data. **Cancer Discovery**, [S. l.], v. 2, n. 5, p. 401–404, 2012.

CHEN, Jeng Ting *et al.* Integrated omics profiling identifies hypoxia-regulated genes in HCT116 colon cancer cells. **Journal of Proteomics**, [S. l.], v. 188, n. 2017, p. 139–151, 2018.

CHIAPPETTA, G. *et al.* High level expression of the HMGI (Y) gene during embryonic development. **Oncogene**, [S. l.], v. 13, n. 11, p. 2439–2446, 1996. Disponível em: <http://europepmc.org/abstract/MED/8957086>

CLAYTON, Aled *et al.* Induction of heat shock proteins in B-cell exosomes. **Journal**

of **Cell Science**, [S. l.], v. 118, n. 16, p. 3631–3638, 2005.

COLAMAIO, Marianna *et al.* HMGA1 silencing reduces stemness and temozolomide resistance in glioblastoma stem cells. **Expert opinion on therapeutic targets**, England, v. 20, n. 10, p. 1169–1179, 2016.

CONWAY, Jake R.; LEX, Alexander; GEHLENBORG, Nils. UpSetR: An R package for the visualization of intersecting sets and their properties. **Bioinformatics**, [S. l.], v. 33, n. 18, p. 2938–2940, 2017.

COX, Thomas R. *et al.* The hypoxic cancer secretome induces pre-metastatic bone lesions through lysyl oxidase. **Nature**, [S. l.], v. 522, n. 7554, p. 106–110, 2015 a.

COX, Thomas R. *et al.* Dataset for the proteomic inventory and quantitative analysis of the breast cancer hypoxic secretome associated with osteotropism. **Data in Brief**, [S. l.], v. 5, p. 621–625, 2015 b. Disponível em: <http://dx.doi.org/10.1016/j.dib.2015.09.039>

CROMPTON, Brian D. *et al.* The genomic landscape of pediatric Ewing sarcoma. **Cancer Discovery**, [S. l.], v. 4, n. 11, p. 1326–1341, 2014.

CRUZ-GARCIA, David *et al.* Remodeling of secretory compartments creates CUPS during nutrient starvation. **The Journal of cell biology**, [S. l.], v. 207, n. 6, p. 695–703, 2014.

CZERWIŃSKA, Patrycja; MAZUREK, Sylwia; WIZNEROWICZ, Maciej. The complexity of TRIM28 contribution to cancer. **Journal of biomedical science**, [S. l.], v. 24, n. 1, p. 63, 2017.

DAYA, D.; NAZERALI, L.; FRANK, G. L. Metastatic ovarian carcinoma of large intestinal origin simulating primary ovarian carcinoma. A clinicopathologic study of 25 cases. **American journal of clinical pathology**, England, v. 97, n. 6, p. 751–758, 1992.

DEUTSCH, Eric W. *et al.* A Guided Tour of the TPP. [S. l.], v. 10, n. 6, p. 1150–1159, 2011.

DEUTSCH, Eric W. *et al.* Trans-Proteomic Pipeline, a standardized data processing pipeline for large-scale reproducible proteomics informatics. **Proteomics - Clinical Applications**, [S. l.], v. 9, n. 7–8, p. 745–754, 2015.

DEUTSCH, Eric W. *et al.* The ProteomeXchange consortium in 2017: Supporting the cultural change in proteomics public data deposition. **Nucleic Acids Research**, [S. l.], v. 45, n. D1, p. D1100–D1106, 2017.

DING, Li *et al.* Perspective on Oncogenic Processes at the End of the Beginning of Cancer Genomics. **Cell**, [S. l.], v. 173, n. 2, p. 305–320.e10, 2018.

EIFLER, Karolin; VERTEGAAL, Alfred C. O. SUMOylation-Mediated Regulation of Cell Cycle Progression and Cancer. **Trends in biochemical sciences**, [S. l.], v. 40, n. 12, p. 779–793, 2015.

ELLROTT, Kyle *et al.* Scalable Open Science Approach for Mutation Calling of Tumor Exomes Using Multiple Genomic Pipelines. **Cell Systems**, [S. l.], v. 6, n. 3, p. 271- 281.e7, 2018.

EMR, Scott *et al.* Journeys through the Golgi--taking stock in a new era. **The Journal of cell biology**, [S. l.], v. 187, n. 4, p. 449–453, 2009.

ENG, Jimmy K. *et al.* A Deeper Look into Comet - Implementation and Features. **Journal of the American Society for Mass Spectrometry**, [S. l.], v. 26, n. 11, p. 1865–1874, 2015.

ENG, Jimmy K.; JAHAN, Tahmina A.; HOOPMANN, Michael R. Comet: An open-source MS/MS sequence database search tool. **Proteomics**, [S. l.], v. 13, n. 1, p. 22–24, 2013.

EUSTACE, Brenda K.; JAY, Daniel G. Extracellular roles for the molecular chaperone, hsp90. **Cell cycle (Georgetown, Tex.)**, [S. l.], v. 3, n. 9, p. 1098—1100, 2004. Disponível em: <http://europepmc.org/abstract/MED/15326368>

EVANS, H. S. *et al.* The risk of subsequent primary cancers after colorectal cancer in southeast England. **Gut**, [S. l.], v. 50, n. 5, p. 647–652, 2002.

FARQUHAR, M. G.; PALADE, G. E. The Golgi apparatus: 100 years of progress and controversy. **Trends in cell biology**, [S. l.], v. 8, n. 1, p. 2–10, 1998.

FEDERICI, Luca; FALINI, Brunangelo. Nucleophosmin mutations in acute myeloid leukemia: a tale of protein unfolding and mislocalization. **Protein science : a publication of the Protein Society**, [S. l.], v. 22, n. 5, p. 545–556, 2013.

FLOTHO, Annette; MELCHIOR, Frauke. Sumoylation: a regulatory protein modification in health and disease. **Annual review of biochemistry**, United States, v. 82, p. 357–385, 2013.

FULDA, Simone *et al.* Cellular stress responses: cell survival and cell death. **International journal of cell biology**, [S. l.], v. 2010, p. 214074, 2010.

GAO, Jianjiong *et al.* Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. **Science Signaling**, [S. l.], v. 6, n. 269, 2013.

GAO, Qingsong *et al.* Driver Fusions and Their Implications in the Development and Treatment of Human Cancers. **Cell Reports**, [S. l.], v. 23, n. 1, p. 227- 238.e3, 2018.

GEE, Heon Yung *et al.* Rescue of $\Delta F508$ -CFTR trafficking via a GRASP-dependent unconventional secretion pathway. **Cell**, United States, v. 146, n. 5, p. 746–760, 2011.

GRISENDI, Silvia *et al.* Role of nucleophosmin in embryonic development and tumorigenesis. **Nature**, England, v. 437, n. 7055, p. 147–153, 2005.

GRØNBORG, Mads *et al.* Biomarker discovery from pancreatic cancer secretome using a differential proteomic approach. **Molecular and Cellular Proteomics**, [S. l.], v. 5, n. 1, p. 157–171, 2006.

GRUNDT, Kirsten *et al.* Characterisation of the NUCKS gene on human chromosome 1q32.1 and the presence of a homologous gene in different species. **Biochemical and biophysical research communications**, United States, v. 323, n. 3, p. 796–801, 2004.

HAJJ, Glaucia N. M. *et al.* The unconventional secretion of stress-inducible protein 1 by a heterogeneous population of extracellular vesicles. **Cellular and Molecular Life Sciences**, [S. l.], v. 70, n. 17, p. 3211–3227, 2013.

HANAHAN, D.; WEINBERG, R. A. The hallmarks of cancer. **Cell**, United States, v. 100, n. 1, p. 57–70, 2000.

HANAHAN, Douglas; WEINBERG, Robert A. Hallmarks of cancer: the next generation. **Cell**, United States, v. 144, n. 5, p. 646–674, 2011.

HAWKINS, Allegra G. *et al.* The Ewing sarcoma secretome and its response to activation of Wnt/beta-catenin signaling. **Molecular and Cellular Proteomics**, [S. l.], v. 17, n. 5, p. 901–912, 2018.

HAY, Ronald T. SUMO: a history of modification. **Molecular cell**, United States, v. 18, n. 1, p. 1–12, 2005.

HERRMANN, Johannes M.; MALKUS, Per; SCHEKMAN, Randy. Out of the ER - Outfitters, escorts and guides. **Trends in Cell Biology**, [S. l.], v. 9, n. 1, p. 5–7, 1999.

HETZER, Martin W.; WENTE, Susan R. Border control at the nucleus: biogenesis and organization of the nuclear membrane and pore complexes. **Developmental cell**, [S. l.], v. 17, n. 5, p. 606–616, 2009.

HILLION, Joelle *et al.* Upregulation of MMP-2 by HMGA1 promotes transformation in undifferentiated, large-cell lung cancer. **Molecular cancer research : MCR**, [S. l.], v. 7, n. 11, p. 1803–1812, 2009.

HOADLEY, Katherine A. *et al.* Cell-of-Origin Patterns Dominate the Molecular Classification of 10,000 Tumors from 33 Types of Cancer. **Cell**, [S. l.], v. 173, n. 2, p. 291–304.e6, 2018.

HOEGE, Carsten *et al.* RAD6-dependent DNA repair is linked to modification of PCNA by ubiquitin and SUMO. **Nature**, England, v. 419, n. 6903, p. 135–141, 2002.

HUANG, Ruixue *et al.* Overexpression of HMGA1 correlates with the malignant status and prognosis of breast cancer. **Molecular and cellular biochemistry**, Netherlands, v. 404, n. 1–2, p. 251–257, 2015.

HÜLSMANN, Bastian B.; LABOKHA, Aksana A.; GÖRLICH, Dirk. The permeability of reconstituted nuclear pores provides direct evidence for the selective phase model. **Cell**, United States, v. 150, n. 4, p. 738–751, 2012.

IRIANTO, Jerome *et al.* Nuclear lamins in cancer. **Cellular and molecular bioengineering**, [S. l.], v. 9, n. 2, p. 258–267, 2016.

JASSAL, Bijay *et al.* The reactome pathway knowledgebase. **Nucleic acids**

research, [S. l.], v. 48, n. D1, p. D498–D503, 2020.

JOHNSTONE, R. M. *et al.* Vesicle formation during reticulocyte maturation. Association of plasma membrane activities with released vesicles (exosomes). **The Journal of biological chemistry**, United States, v. 262, n. 19, p. 9412–9420, 1987.

KÄLL, Lukas; KROGH, Anders; SONNHAMMER, Erik L. L. Advantages of combined transmembrane topology and signal peptide prediction--the Phobius web server. **Nucleic acids research**, [S. l.], v. 35, n. Web Server issue, p. W429-32, 2007.

KALLURI, Raghu; LEBLEU, Valerie S. The biology, function, and biomedical applications of exosomes. **Science (New York, N.Y.)**, [S. l.], v. 367, n. 6478, 2020.

KEERTHIKUMAR, Shivakumar *et al.* ExoCarta: A Web-Based Compendium of Exosomal Cargo. **Journal of Molecular Biology**, [S. l.], v. 428, n. 4, p. 688–692, 2016. Disponível em: <http://dx.doi.org/10.1016/j.jmb.2015.09.019>

KELLER, Andrew *et al.* Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. **Analytical Chemistry**, [S. l.], v. 74, n. 20, p. 5383–5392, 2002.

KESSENBROCK, Kai; PLAKS, Vicki; WERB, Zena. Matrix metalloproteinases: regulators of the tumor microenvironment. **Cell**, [S. l.], v. 141, n. 1, p. 52–67, 2010.

KETTUNEN, Eeva *et al.* Differentially expressed genes in nonsmall cell lung cancer: expression profiling of cancer-related genes in squamous cell lung cancer. **Cancer genetics and cytogenetics**, United States, v. 149, n. 2, p. 98–106, 2004.

KIM, Jiyeon; GEE, Heon Yung; LEE, Min Goo. Unconventional protein secretion – new insights into the pathogenesis and therapeutic targets of human diseases. **Journal of Cell Science**, [S. l.], v. 131, n. 12, p. jcs213686, 2018.

KLUNE, John R. *et al.* **HMGB1: Endogenous danger signaling**. [S. l.: s. n.]

KÖHLER, Alwin; HURT, Ed. Exporting RNA from the nucleus to the cytoplasm. **Nature reviews. Molecular cell biology**, England, v. 8, n. 10, p. 761–773, 2007.

LA VENUTA, Giuseppe *et al.* The Startling Properties of Fibroblast Growth Factor 2: How to Exit Mammalian Cells without a Signal Peptide at Hand. **The Journal of biological chemistry**, [S. l.], v. 290, n. 45, p. 27015–27020, 2015.

LADINSKY, M. S. *et al.* Golgi structure in three dimensions: functional insights from the normal rat kidney cell. **The Journal of cell biology**, [S. l.], v. 144, n. 6, p. 1135–1149, 1999.

LANCASTER, Graeme I.; FEBBRAIO, Mark A. Exosome-dependent trafficking of HSP70: A novel secretory pathway for cellular stress proteins. **Journal of Biological Chemistry**, [S. l.], v. 280, n. 24, p. 23349–23355, 2005.

LAWLOR, Kevin *et al.* Pathway-based biomarker search by high-throughput proteomics profiling of secretomes. **Journal of Proteome Research**, [S. l.], v. 8, n. 3, p. 1489–1503, 2009.

LEX, Alexander *et al.* UpSet: Visualization of Intersecting Sets Europe PMC Funders Group. **IEEE Trans Vis Comput Graph**, [S. l.], v. 20, n. 12, p. 1983–1992, 2014. Disponível em: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4720993/pdf/emss-66749.pdf>

LIBERATO, Tarcísio *et al.* Signatures of protein expression revealed by secretome analyses of cancer associated fibroblasts and melanoma cell lines. **Journal of Proteomics**, [S. l.], v. 174, n. December 2017, p. 1–8, 2018.

LIN, Qifeng *et al.* Sieving through the cancer secretome. **Biochimica et biophysica acta**, Netherlands, v. 1834, n. 11, p. 2360–2371, 2013.

LIU, Jianfang *et al.* An Integrated TCGA Pan-Cancer Clinical Data Resource to Drive High-Quality Survival Outcome Analytics. **Cell**, [S. l.], v. 173, n. 2, p. 400–416.e11, 2018.

LONGSHAW, Victoria M. *et al.* Nuclear translocation of the Hsp70/Hsp90 organizing protein mSTI1 is regulated by cell cycle kinases. **Journal of Cell Science**, [S. l.], v. 117, n. 5, p. 701–710, 2004.

LOPES, Marilene H. *et al.* Interaction of cellular prion and stress-inducible protein 1 promotes neuritogenesis and neuroprotection by distinct signaling pathways. **Journal of Neuroscience**, [S. l.], v. 25, n. 49, p. 11330–11339, 2005.

MA, Chao *et al.* CBX3 predicts an unfavorable prognosis and promotes tumorigenesis in osteosarcoma. **Molecular medicine reports**, [S. l.], v. 19, n. 5, p. 4205–4212, 2019.

MA, Xiao Lu *et al.* Serum STIP1, a Novel Indicator for Microvascular Invasion, Predicts Outcomes and Treatment Response in Hepatocellular Carcinoma. **Frontiers in Oncology**, [S. l.], v. 10, n. April, p. 1–12, 2020.

MACIEJEWSKI, Andrzej *et al.* Domains of STIP1 responsible for regulating PrPC-dependent amyloid- β oligomer toxicity. **Biochemical Journal**, [S. l.], v. 473, n. 14, p. 2119–2130, 2016.

MACKENZIE, Amanda *et al.* Rapid secretion of interleukin-1 β by microvesicle shedding. **Immunity**, [S. l.], v. 15, n. 5, p. 825–835, 2001.

MARCEAU, Aimee H. Functions of single-strand DNA-binding proteins in DNA replication, recombination, and repair. **Methods in molecular biology (Clifton, N.J.)**, United States, v. 922, p. 1–21, 2012.

MARTINS, V. R. *et al.* Stress-inducible protein 1 is a cell surface ligand for cellular prion that triggers neuroprotection. **Embo Journal**, [S. l.], v. 21, n. 13, p. 3307–3316, 2002.

MATHIAS, Rommel A. *et al.* Secretome-based proteomic profiling of Ras-transformed MDCK cells reveals extracellular modulators of epithelial-mesenchymal transition. **Journal of Proteome Research**, [S. l.], v. 8, n. 6, p. 2827–2837, 2009.

MATUNIS, M. J.; COUTAVAS, E.; BLOBEL, G. A novel ubiquitin-like modification modulates the partitioning of the Ran-GTPase-activating protein RanGAP1 between the cytosol and the nuclear pore complex. **The Journal of cell biology**, [S. l.], v. 135, n. 6 Pt 1, p. 1457–1470, 1996.

MENDEZ, Olga *et al.* Extracellular HMGA1 promotes tumor invasion and metastasis in triple-negative breast cancer. **Clinical Cancer Research**, [S. l.], v. 24, n. 24, p. 6367–6382, 2018.

MÉNDEZ, Olga *et al.* **Clinical implications of extracellular HMGA1 in breast cancer**. [S. l.: s. n.]

MIN, Kyung Won; LEE, Seong Ho; BAEK, Seung Joon. Moonlighting proteins in cancer. **Cancer Letters**, [S. l.], v. 370, n. 1, p. 108–116, 2016. Disponível em: <http://dx.doi.org/10.1016/j.canlet.2015.09.022>

MORTON, Lindsay M. *et al.* The rising incidence of second cancers: patterns of occurrence and identification of risk factors for children and adults. **American Society of Clinical Oncology educational book. American Society of Clinical Oncology. Annual Meeting**, United States, p. e57-67, 2014.

NEO, Shu Hui *et al.* TRIM28 Is an E3 Ligase for ARF-Mediated NPM1/B23 SUMOylation That Represses Centrosome Amplification. **Molecular and Cellular Biology**, [S. l.], v. 35, n. 16, p. 2851–2863, 2015.

NICKEL, Walter; RABOUILLE, Catherine. Mechanisms of regulated unconventional protein secretion Walter. **Nature Reviews Molecular Cell Biology**, [S. l.], v. 10, n. February, p. 148–155, 2009. Disponível em: <papers://01628a62-bdc0-47f5-904e-9689163247b2/Paper/p72>

NISHIMURA, Yuki *et al.* Tryptophans 286 and 288 in the C-terminal region of protein B23.1 are important for its nucleolar localization. **Bioscience, biotechnology, and biochemistry**, England, v. 66, n. 10, p. 2239–2242, 2002.

ODUNUGA, O. O.; LONGSHAW, V. M.; BLATCH, G. L. Hop: More than an Hsp70/Hsp90 adaptor protein. **BioEssays**, [S. l.], v. 26, n. 10, p. 1058–1068, 2004.

OSTAPCHENKO, Valeriy G. *et al.* The prion protein ligand, stress-inducible phosphoprotein 1, regulates amyloid- β oligomer toxicity. **Journal of Neuroscience**, [S. l.], v. 33, n. 42, p. 16552–16564, 2013.

OSTVOLD, A. C.; HOLTLUND, J.; LALAND, S. G. A novel, highly phosphorylated protein, of the high-mobility group type, present in a variety of proliferating and non-proliferating mammalian cells. **European journal of biochemistry**, England, v. 153, n. 3, p. 469–475, 1985.

OTT, Martin; AMUNTS, Alexey; BROWN, Alan. Organization and Regulation of Mitochondrial Protein Synthesis. **Annual review of biochemistry**, United States, v. 85, p. 77–101, 2016.

PALADE, G. Intracellular aspects of the process of protein synthesis. **Science (New**

York, N.Y.), United States, v. 189, n. 4200, p. 347–358, 1975.

PALTRIDGE, James L.; BELLE, Leila; KHEW-GOODALL, Yeesim. The secretome in cancer progression. **Biochimica et Biophysica Acta - Proteins and Proteomics**, [S. l.], v. 1834, n. 11, p. 2233–2241, 2013. Disponível em: <http://dx.doi.org/10.1016/j.bbapap.2013.03.014>

PAPAEMMANUIL, Elli *et al.* Genomic Classification and Prognosis in Acute Myeloid Leukemia. **The New England journal of medicine**, [S. l.], v. 374, n. 23, p. 2209–2221, 2016.

PARK, Ji-Hyeon; JUNG, Dong Hae; BAEK, Jeong-Heum. Metastatic colon cancer of an ovarian cancer origin mimicking primary colon cancer: A case report. **Korean Journal of Clinical Oncology**, [S. l.], v. 14, n. 1, p. 53–57, 2018.

PATEL, Anoop P. *et al.* Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. **Science (New York, N.Y.)**, [S. l.], v. 344, n. 6190, p. 1396–1401, 2014.

PATHAN, Mohashin *et al.* Vesiclepedia 2019: A compendium of RNA, proteins, lipids and metabolites in extracellular vesicles. **Nucleic Acids Research**, [S. l.], v. 47, n. D1, p. D516–D519, 2019.

PFEFFER, Suzanne R. How the Golgi works: a cisternal progenitor model. **Proceedings of the National Academy of Sciences of the United States of America**, [S. l.], v. 107, n. 46, p. 19614–19618, 2010.

PICHLER, Andrea *et al.* The nucleoporin RanBP2 has SUMO1 E3 ligase activity. **Cell**, United States, v. 108, n. 1, p. 109–120, 2002.

PUHKA, Maija *et al.* Metabolomic Profiling of Extracellular Vesicles and Alternative Normalization Methods Reveal Enriched Metabolites and Strategies to Study Prostate Cancer-Related Changes. **Theranostics**, [S. l.], v. 7, n. 16, p. 3824–3841, 2017.

PUTZ, Ulrich *et al.* The tumor suppressor PTEN is exported in exosomes and has phosphatase activity in recipient cells. **Science Signaling**, [S. l.], v. 5, n. 243, 2012.

QU, Yan *et al.* Nonclassical IL-1 β Secretion Stimulated by P2X7 Receptors Is Dependent on Inflammasome Activation and Correlated with Exosome Release in Murine Macrophages. **The Journal of Immunology**, [S. l.], v. 179, n. 3, p. 1913–1925, 2007.

QUIMBY, B. B.; CORBETT, A. H. Nuclear transport mechanisms. **Cellular and molecular life sciences : CMLS**, Switzerland, v. 58, n. 12–13, p. 1766–1773, 2001.

RABELLINO, Andrea; ANDREANI, Cristina; SCAGLIONI, Pier Paolo. The Role of PIAS SUMO E3-Ligases in Cancer. **Cancer research**, [S. l.], v. 77, n. 7, p. 1542–1547, 2017.

RABOUILLE, Catherine. Pathways of Unconventional Protein Secretion. **Trends in**

Cell Biology, [S. l.], v. 27, n. 3, p. 230–240, 2017. Disponível em: <http://dx.doi.org/10.1016/j.tcb.2016.11.007>

RADISKY, Derek C. *et al.* Single proteins might have dual but related functions in intracellular and extracellular microenvironments. **Nature Reviews Molecular Cell Biology**, [S. l.], v. 10, n. 3, p. 228–234, 2009.

RODRÍGUEZ, Luz-Stella *et al.* Immunomodulators released during rotavirus infection of polarized caco-2 cells. **Viral immunology**, United States, v. 22, n. 3, p. 163–172, 2009.

ROFFÉ, Martín *et al.* Prion protein interaction with stress-inducible protein 1 enhances neuronal protein synthesis via mTOR. **Proceedings of the National Academy of Sciences of the United States of America**, [S. l.], v. 107, n. 29, p. 13147–13152, 2010.

ROTHBALLER, Andrea; KUTAY, Ulrike. Poring over pores: nuclear pore complex insertion into the nuclear envelope. **Trends in biochemical sciences**, England, v. 38, n. 6, p. 292–301, 2013.

RUIVO, Carolina F. *et al.* The Biology of Cancer Exosomes: Insights and New Perspectives. **Cancer research**, United States, v. 77, n. 23, p. 6480–6488, 2017.

RUSSO, Jose *et al.* The genomic signature of breast cancer prevention. **Recent results in cancer research. Fortschritte der Krebsforschung. Progres dans les recherches sur le cancer**, Germany, v. 174, p. 131–150, 2007.

SANCHEZ-VEGA, Francisco *et al.* Oncogenic Signaling Pathways in The Cancer Genome Atlas. **Cell**, [S. l.], v. 173, n. 2, p. 321–337.e10, 2018.

SANTIAGO, Aleixo *et al.* P53 SUMOylation promotes its nuclear export by facilitating its release from the nuclear export receptor CRM1. **Molecular Biology of the Cell**, [S. l.], v. 24, n. 17, p. 2739–2752, 2013.

SANTOS, Tiago Góss; MARTINS, Vilma Regina; HAJJ, Glaucia Noeli Maroso. Unconventional Secretion of Heat Shock Proteins in Cancer. **International journal of molecular sciences**, [S. l.], v. 18, n. 5, 2017.

SCAFFIDI, Paola; MISTELI, Tom; BIANCHI, Marco E. Release of chromatin protein HMGB1 by necrotic cells triggers inflammation. **Nature**, [S. l.], v. 418, n. 6894, p. 191–195, 2002. Disponível em: <https://doi.org/10.1038/nature00858>

SCHAAIJ-VISSER, Tienke B. M. *et al.* The cancer secretome, current status and opportunities in the lung, breast and colorectal cancer context. **Biochimica et biophysica acta**, Netherlands, v. 1834, n. 11, p. 2242–2258, 2013.

SCHEY, Kevin L.; LUTHER, J. Matthew; ROSE, Kristie L. Proteomics characterization of exosome cargo. **Methods (San Diego, Calif.)**, [S. l.], v. 87, p. 75–82, 2015.

SCHOTMAN, Hans; KARHINEN, Leena; RABOUILLE, Catherine. dGRASP-mediated

noncanonical integrin secretion is required for *Drosophila* epithelial remodeling. **Developmental cell**, United States, v. 14, n. 2, p. 171–182, 2008.

SCHRODER, Kate; TSCHOPP, Jurg. The inflammasomes. **Cell**, United States, v. 140, n. 6, p. 821–832, 2010.

SEELER, Jacob-Sebastian; DEJEAN, Anne. SUMO and the robustness of cancer. **Nature reviews. Cancer**, England, v. 17, n. 3, p. 184–197, 2017.

SHUCH, BRIAN; LINEHAN, B. W. M.L.; SRIVASAN, R. Molecular Characterization of Basal-Like and Non-Basal-Like Triple-Negative Breast Cancer. *[S. l.]*, v. 0031, n. 0, p. 1051–1062, 2012.

SKOTLAND, Tore; SANDVIG, Kirsten; LLORENTE, Alicia. Lipids in exosomes: Current knowledge and the way forward. **Progress in lipid research**, England, v. 66, p. 30–41, 2017.

STERINGER, Julia P.; MÜLLER, Hans-Michael; NICKEL, Walter. Unconventional secretion of fibroblast growth factor 2—a novel type of protein translocation across membranes? **Journal of molecular biology**, England, v. 427, n. 6 Pt A, p. 1202–1210, 2015.

STERINGER, Julia P.; NICKEL, Walter. A direct gateway into the extracellular space: Unconventional secretion of FGF2 through self-sustained plasma membrane pores. **Seminars in cell & developmental biology**, England, v. 83, p. 3–7, 2018.

STRAMBIO-DE-CASTILLIA, Caterina; NIEPEL, Mario; ROUT, Michael P. The nuclear pore complex: bridging nuclear transport and gene regulation. **Nature reviews. Molecular cell biology**, England, v. 11, n. 7, p. 490–501, 2010.

TAI, Yu-Ling *et al.* Exosomes in cancer development and clinical applications. **Cancer science**, *[S. l.]*, v. 109, n. 8, p. 2364–2374, 2018.

TAIT, Stephen W. G.; GREEN, Douglas R. Mitochondria and cell death: Outer membrane permeabilization and beyond. **Nature Reviews Molecular Cell Biology**, *[S. l.]*, v. 11, n. 9, p. 621–632, 2010. Disponível em: <http://dx.doi.org/10.1038/nrm2952>

TAYLOR, Alison M. *et al.* Genomic and Functional Approaches to Understanding Cancer Aneuploidy. **Cancer Cell**, *[S. l.]*, v. 33, n. 4, p. 676–689.e3, 2018.

THUL, Peter J. *et al.* A subcellular map of the human proteome. **Science (New York, N.Y.)**, United States, v. 356, n. 6340, 2017.

TJALSMA, H. *et al.* Signal peptide-dependent protein transport in *Bacillus subtilis*: a genome-based survey of the secretome. **Microbiology and molecular biology reviews : MMBR**, *[S. l.]*, v. 64, n. 3, p. 515–547, 2000.

TORRADO, Lucía Cespón *et al.* An intrinsic quality-control mechanism ensures unconventional secretion of fibroblast growth factor 2 in a folded conformation. **Journal of cell science**, England, v. 122, n. Pt 18, p. 3322–3329, 2009.

TRAN, Elizabeth J.; WENTE, Susan R. Dynamic nuclear pore complexes: life on the edge. **Cell**, United States, v. 125, n. 6, p. 1041–1053, 2006.

TSAI, Chia Lung *et al.* Secreted Stress-Induced Phosphoprotein 1 Activates the ALK2-SMAD Signaling Pathways and Promotes Cell Proliferation of Ovarian Cancer Cells. **Cell Reports**, [S. l.], v. 2, n. 2, p. 283–293, 2012. Disponível em: <http://dx.doi.org/10.1016/j.celrep.2012.07.002>

UHLÉN, Mathias *et al.* A pathology atlas of the human cancer transcriptome. **Science (New York, N.Y.)**, United States, v. 357, n. 6352, 2017.

UHLÉN, Mathias *et al.* Tissue-based map of the human proteome. **Science**, [S. l.], v. 347, n. 6220, 2015.

UHLÉN, Mathias *et al.* The human secretome. **Science Signaling**, [S. l.], v. 12, n. 609, p. 1–9, 2019.

VAN DEN BOORN, Jasper G. *et al.* Exosomes as nucleic acid nanocarriers. **Advanced drug delivery reviews**, Netherlands, v. 65, n. 3, p. 331–335, 2013.

VARKE, Ajit. Evolutionary forces shaping the Golgi glycosylation machinery: why cell surface glycans are universal to living cells. **Cold Spring Harbor perspectives in biology**, [S. l.], v. 3, n. 6, 2011.

VIKLUND, Håkan *et al.* SPOCTOPUS: a combined predictor of signal peptides and membrane protein topology. **Bioinformatics (Oxford, England)**, England, v. 24, n. 24, p. 2928–2929, 2008.

VILLARREAL, Laura *et al.* Unconventional secretion is a major contributor of cancer cell line secretomes. **Molecular and Cellular Proteomics**, [S. l.], v. 12, n. 5, p. 1046–1060, 2013.

VIOTTI, Corrado. ER to Golgi-Dependent Protein Secretion: The Conventional Pathway. **Methods in molecular biology (Clifton, N.J.)**, United States, v. 1459, p. 3–29, 2016.

VIZCAÍNO, Juan A. *et al.* ProteomeXchange provides globally coordinated proteomics data submission and dissemination. **Nature Biotechnology**, [S. l.], v. 32, n. 3, p. 223–226, 2014.

VON HEIJNE, G. The signal peptide. **The Journal of membrane biology**, United States, v. 115, n. 3, p. 195–201, 1990.

VON HEIJNE, Gunnar. **Introduction to theme “membrane protein folding and insertion”**. United States [s. n.]

WALTER, Peter; RON, David. The unfolded protein response: from stress pathway to homeostatic regulation. **Science (New York, N.Y.)**, United States, v. 334, n. 6059, p. 1081–1086, 2011.

WANG, Tzu Hao *et al.* Stress-induced phosphoprotein 1 as a secreted biomarker for human ovarian cancer promotes cancer cell proliferation. **Molecular and Cellular**

Proteomics, [S. l.], v. 9, n. 9, p. 1873–1884, 2010.

WEGEHINGEL, Sabine; ZEHE, Christoph; NICKEL, Walter. Rerouting of fibroblast growth factor 2 to the classical secretory pathway results in post-translational modifications that block binding to heparan sulfate proteoglycans. **FEBS letters**, England, v. 582, n. 16, p. 2387–2392, 2008.

WEINSTEIN, John N. *et al.* The cancer genome atlas pan-cancer analysis project. **Nature Genetics**, [S. l.], v. 45, n. 10, p. 1113–1120, 2013. Disponível em: <http://dx.doi.org/10.1038/ng.2764>

WILLIAMS, Michael D. *et al.* HMGA1 drives metabolic reprogramming of intestinal epithelium during hyperproliferation, polyposis, and colorectal carcinogenesis. **Journal of proteome research**, United States, v. 14, n. 3, p. 1420–1431, 2015.

WORZFELD, Thomas *et al.* Proteotranscriptomics Reveal Signaling Networks in the Ovarian Cancer Microenvironment. **Molecular and Cellular Proteomics**, [S. l.], v. 17, n. 2, p. 270–289, 2018.

XU, Guoxing *et al.* MiR-142-3p functions as a potential tumor suppressor in human osteosarcoma by targeting HMGA1. **Cellular physiology and biochemistry : international journal of experimental cellular physiology, biochemistry, and pharmacology**, Germany, v. 33, n. 5, p. 1329–1339, 2014.

XUE, Hua; LU, Bingjian; LAI, Maode. The cancer secretome: A reservoir of biomarkers. **Journal of Translational Medicine**, [S. l.], v. 6, p. 1–12, 2008.

YUAN, Zheng *et al.* Computational differentiation of N-terminal signal peptides and transmembrane helices. **Biochemical and biophysical research communications**, United States, v. 312, n. 4, p. 1278–1283, 2003.

ZARKA, Jabra *et al.* Nucleophosmin 1 Mutations in Acute Myeloid Leukemia. **Genes (Basel)**, [S. l.], v. 11, n. 6, p. 1–16, 2020.

ZHAN, Xiechao *et al.* Structures of the human pre-catalytic spliceosome and its precursor spliceosome. **Cell research**, [S. l.], v. 28, n. 12, p. 1129–1140, 2018.

ZHANG, Xiaofeng *et al.* Structure of the human activated spliceosome in three conformational states. **Cell research**, [S. l.], v. 28, n. 3, p. 307–322, 2018.

ANEXOS

ANEXO A – Visão geral dos arquivos selecionados e ferramentas utilizadas pelos respectivos autores nas buscas

Tipo de câncer	ID dos arquivos	Quantidade de arquivos (tamanho total)	Publicação associada	Instrumento	Software	Banco de dados	Número de proteínas identificadas (FDR ≤ 1%)
Cólon	PXD009009	44 (8,38 GB)	(CHEN <i>et al.</i> , 2018)	LTQ-Orbitrap (Thermo Fisher Scientific)	Proteome Discoverer 1.4.0.288/Mascot 2.2	sprot_2014.fasta (20.266 sequências)	1740
Mama	PXD000397	4 (9,15 GB)	(COX <i>et al.</i> , 2015a, 2015b)	QExactive (ThermoFisher Scientific)	MaxQuant 1.5.0.0/Perseus 1.4	Human EnsemblGRCh38	1426
Melanoma	PXD007265	6 (4,49 GB)	(LIBERATO <i>et al.</i> , 2018)	LTQ-Orbitrap Velos (Thermo Fisher Scientific)	MaxQuant 1.5.5.1	UniProt 2017_07 Taxonomia: <i>Homo sapiens</i> (159.552 sequências)	A375: 1020 HS895T: 662 SH4: 1182 (listas unificadas)
Ovário	PXD008047	40 (35,9 GB)	(WORZFELD <i>et al.</i> , 2018)	QExactive HF (ThermoFisher Scientific)	MaxQuant 1.5.5.1	Ensembl 81 (101.933 sequências)	5416
Sarcoma de Ewing	PXD007909	6 (10,9 GB)	(HAWKINS <i>et al.</i> , 2018)	QExactive HF (ThermoFisher Scientific)	Proteome Discoverer 2.1	SwissProt v2016-11-30 (20.213 sequências)	CHLA10: 1900 TC32: 2270 (listas unificadas)

ANEXO B – Características específicas das amostras de câncer de cólon selecionadas

Características da amostra	
Linhagem celular	HCT116 (ATCC® CCL-247™)
Características	Carcinoma colorretal

Parâmetros de busca	
Equipamento	LTQ-Orbitrap (Thermo Fisher Scientific)
Método (MS)	High-Low
Banco de dados	sprot_2014.fasta (Homo sapiens)
Enzima/clivagens perdidas	tripsina/2 clivagens perdidas
Modificações fixas	carbamidometilação (cisteína)
Modificações variáveis	Dimethyl: 2H(4)13C(2) (K), Dimethyl: 2H(4)13C(2) (N-term), Dimethyl (K) and Dimethyl (N-term)

(continua)

Arquivos brutos escolhidos		
Descrição	Nome	Tamanho
Dados do secretoma na condição normoxia ^H /hipóxia ^L (Primeiro experimento) Pesado: Normoxia (21% O ₂)	20151211-hypoxia-12_5ug-s01.RAW	132,772 MB
	20151211-hypoxia-12_5ug-s02.RAW	132,703 MB
	20151211-hypoxia-12_5ug-s03.RAW	198,351 MB
	20151211-hypoxia-12_5ug-s04.RAW	117,131 MB
	20151211-hypoxia-12_5ug-s05.RAW	132,275 MB
	20151211-hypoxia-12_5ug-s06.RAW	154,811 MB
	20151211-hypoxia-12_5ug-s07.RAW	117,741 MB
	20151211-hypoxia-12_5ug-s08.RAW	142,135 MB
	20151211-hypoxia-12_5ug-s09.RAW	158,051 MB
	20151211-hypoxia-12_5ug-s10.RAW	191,392 MB
	20151211-hypoxia-12_5ug-s11.RAW	136,384 MB
	20151211-hypoxia-12_5ug-s12.RAW	158,351 MB
	20151211-hypoxia-12_5ug-s13.RAW	154,622 MB
	20151211-hypoxia-12_5ug-s14.RAW	214,384 MB
	20151211-hypoxia-12_5ug-s15.RAW	181,267 MB
	20151211-hypoxia-12_5ug-s16.RAW	265,843 MB
	20151211-hypoxia-12_5ug-s17.RAW	190,982 MB
	20151211-hypoxia-12_5ug-s18.RAW	274,23 MB
	20151211-hypoxia-12_5ug-s19.RAW	225,029 MB
	20151211-hypoxia-12_5ug-s20.RAW	309,497 MB
	20151211-hypoxia-12_5ug-s21.RAW	279,796 MB
	20151211-hypoxia-12_5ug-s22.RAW	229,901 MB

Arquivos brutos escolhidos		
Descrição	Nome	Tamanho
Dados do secretoma na condição hipóxia ^H /normoxia ^L (Segundo experimento) Leve: Normoxia (21% O ₂)	20151214-hypoxia-12_5ug-s01.RAW	189,849 MB
	20151214-hypoxia-12_5ug-s02.RAW	197,8 MB
	20151214-hypoxia-12_5ug-s03.RAW	186,535 MB
	20151214-hypoxia-12_5ug-s04.RAW	162,585 MB
	20151214-hypoxia-12_5ug-s05.RAW	141,405 MB
	20151214-hypoxia-12_5ug-s06.RAW	148,771 MB
	20151214-hypoxia-12_5ug-s07.RAW	141,38 MB
	20151214-hypoxia-12_5ug-s08.RAW	196,371 MB
	20151214-hypoxia-12_5ug-s09.RAW	198,516 MB
	20151214-hypoxia-12_5ug-s10.RAW	216,409 MB
	20151214-hypoxia-12_5ug-s11.RAW	155,548 MB
	20151214-hypoxia-12_5ug-s12.RAW	168,148 MB
	20151214-hypoxia-12_5ug-s13.RAW	165,306 MB
	20151214-hypoxia-12_5ug-s14.RAW	230,054 MB
	20151214-hypoxia-12_5ug-s15.RAW	205,597 MB
	20151214-hypoxia-12_5ug-s16.RAW	267,033 MB
	20151214-hypoxia-12_5ug-s17.RAW	206,297 MB
	20151214-hypoxia-12_5ug-s18.RAW	268,497 MB
	20151214-hypoxia-12_5ug-s19.RAW	233,762 MB
	20151214-hypoxia-12_5ug-s20.RAW	301,082 MB
	20151214-hypoxia-12_5ug-s21.RAW	286,033 MB
	20151214-hypoxia-12_5ug-s22.RAW	225,565 MB

ANEXO C – Características específicas das amostras de câncer de mama selecionadas

Características da amostra	
Linhagem celular	MDAMB231 (ATCC® HTB-26™)
Características	Tecido da glândula mamária derivada do sítio metastático efusão pleural

Parâmetros de busca	
Equipamento	QExactive (ThermoFisher Scientific)
Método (MS)	High-high
Banco de dados	EnsemblGRCh38 (Homo sapiens)
Enzima/clivagens perdidas	tripsina/2 clivagens perdidas
Modificações fixas	carbamidometilação (cisteína)
Modificações variáveis	oxidação (metionina), acetilação (N-terminal)

Arquivos brutos escolhidos		
Descrição	Nome	Tamanho
Replicata técnica 1	MDA_wt_Air.raw	2.056,367 MB
Replicata técnica 2	MDA_wt_Air_rep2.raw	1.972,382 MB

ANEXO D – Características específicas das amostras de melanoma selecionadas

Características das amostras	
Linhagem celular	Características
HS895T (ATCC® CRL-7637™)	Fibroblasto associado ao câncer
A375 (ATCC® CRL-1619™)	Melanócito tumoral metastático
SH4 (ATCC® CRL-7724™)	Célula extraída de efusão pleural de paciente com melanoma

Parâmetros de busca	
Equipamento	LTQ-Orbitrap Velos (Thermo Fisher Scientific)
Método (MS)	High-Low
Database	UniProt 2017_07 (Homo sapiens)
Enzima/clivagens perdidas	tripsina/2 clivagens perdidas
Modificações fixas	carbamidometilação (cisteína); dimetilação (leve/pesado, delta 8)
Modificações variáveis	oxidação (metionina), acetilação (N-terminal), desamidação (glutamina/asparagina)

Arquivos brutos escolhidos		
Descrição	Nome	Tamanho
Replicata biológica 1 Pesado: HS95T	OB_AZP_1_20160827_01.raw	825,94 MB
Replicata biológica 2 Pesado: HS895T	OB_AZP_2_20160827_02.raw	770,525 MB
Replicata biológica 3 Pesado: HS895T	OB_AZP_3_20160827_03.raw	795,262 MB
Replicata biológica 1 A375 ^L /SH4 ^H	OB_AZP_4_20160827_04_160905133413.raw	743,027 MB
Replicata biológica 2 A375 ^L /SH4 ^H	OB_AZP_5_20160827_05_160905163243.raw	739,551 MB
Replicata biológica 3 A375 ^L /SH4 ^H	OB_AZP_6_20160827_06_160905193228.raw	732,89 MB

ANEXO E – Características específicas das amostras de câncer de ovário selecionadas

Característica da amostra	
Subtipo do câncer ovariano estudado	Adenocarcinoma seroso de alto grau
Amostra	Células do subtipo de câncer ovariano adenocarcinoma seroso de alto grau (histologicamente verificado), separadas através da coleta do líquido ascítico de um paciente pós-cirúrgico cultivadas em cultura para a obtenção do secretoma

Parâmetros de busca	
Equipamento	QExactive HF (ThermoFisher Scientific)
Método (MS)	High-high
Database	Ensembl 81 (Homo sapiens)
Enzima/clivagens perdidas	tripsina/2 clivagens perdidas
Modificações fixas	carbamidometilação (cisteína)
Modificações variáveis	oxidação (metionina); acetilação (N-terminal)

(continua)

Arquivos brutos escolhidos		
Descrição	Nome	Tamanho
Secretoma das células esferoides tumorais de um paciente	20170110_SJ_ThomasWorzfeld_SN_153_Tumor_01.raw	1.002,045 MB
	20170110_SJ_ThomasWorzfeld_SN_153_Tumor_02.raw	863,361 MB
	20170110_SJ_ThomasWorzfeld_SN_153_Tumor_03.raw	884,667 MB
	20170110_SJ_ThomasWorzfeld_SN_153_Tumor_04.raw	905,806 MB
	20170110_SJ_ThomasWorzfeld_SN_153_Tumor_05.raw	835,433 MB
	20170110_SJ_ThomasWorzfeld_SN_153_Tumor_06.raw	893,814 MB
	20170110_SJ_ThomasWorzfeld_SN_153_Tumor_07.raw	958,858 MB
	20170110_SJ_ThomasWorzfeld_SN_153_Tumor_08.raw	983,16 MB
	20170110_SJ_ThomasWorzfeld_SN_153_Tumor_09.raw	1.093,837 MB
	20170110_SJ_ThomasWorzfeld_SN_153_Tumor_10.raw	1.067,321 MB
	20170110_SJ_ThomasWorzfeld_SN_159_Tumor_31.raw	1.120,465 MB
	20170110_SJ_ThomasWorzfeld_SN_159_Tumor_32.raw	1.045,719 MB
	20170110_SJ_ThomasWorzfeld_SN_159_Tumor_33.raw	1.086,736 MB
	20170110_SJ_ThomasWorzfeld_SN_159_Tumor_34.raw	1.025,66 MB
	20170110_SJ_ThomasWorzfeld_SN_159_Tumor_35.raw	935,633 MB
	20170110_SJ_ThomasWorzfeld_SN_159_Tumor_36.raw	967,97 MB
	20170110_SJ_ThomasWorzfeld_SN_159_Tumor_37.raw	1.028,733 MB
	20170110_SJ_ThomasWorzfeld_SN_159_Tumor_38.raw	1.077,529 MB
	20170110_SJ_ThomasWorzfeld_SN_159_Tumor_39.raw	1.091,36 MB
	20170110_SJ_ThomasWorzfeld_SN_159_Tumor_40.raw	1.126,635 MB
	20170127_AK_Th_Worzfeld_SN_P161_Tumor_01.raw	826,992 MB
	20170127_AK_Th_Worzfeld_SN_P161_Tumor_02.raw	821,234 MB
	20170127_AK_Th_Worzfeld_SN_P161_Tumor_03.raw	798,582 MB

(conclusão)

Arquivos brutos escolhidos		
Descrição	Nome	Tamanho
Secretoma das células esferoides tumorais de um paciente	20170127_AK_Th_Worzfeld_SN_P161_Tumor_08.raw	803,35 MB
	20170127_AK_Th_Worzfeld_SN_P161_Tumor_09.raw	822,684 MB
	20170127_AK_Th_Worzfeld_SN_P161_Tumor_10.raw	880,591 MB
	20170127_AK_Th_Worzfeld_SN_P162_Tumor_31.raw	881,512 MB
	20170127_AK_Th_Worzfeld_SN_P162_Tumor_32.raw	885,045 MB
	20170127_AK_Th_Worzfeld_SN_P162_Tumor_33.raw	858,297 MB
	20170127_AK_Th_Worzfeld_SN_P162_Tumor_34.raw	851,687 MB
	20170127_AK_Th_Worzfeld_SN_P162_Tumor_35.raw	774,1 MB
	20170127_AK_Th_Worzfeld_SN_P162_Tumor_36.raw	866,42 MB
	20170127_AK_Th_Worzfeld_SN_P162_Tumor_37.raw	907,06 MB
	20170127_AK_Th_Worzfeld_SN_P162_Tumor_38.raw	833,444 MB
	20170127_AK_Th_Worzfeld_SN_P162_Tumor_39.raw	908,942 MB
	20170127_AK_Th_Worzfeld_SN_P162_Tumor_40.raw	914,847 MB

ANEXO F – Características específicas das amostras de sarcoma de Ewing selecionadas

Características das amostras	
Linhagem celular	Características
CHLA-10 (RRID:CVCL_6583)	Tumor neuroectodérmico primitivo
TC-32 (RRID:CVCL_7151)	Tumor neuroectodérmico primitivo

Parâmetros de busca	
Equipamento	QExactive HF (ThermoFisher Scientific)
Método (MS)	High-High
Database	Swissprot v2016-11-30 (Homo sapiens)
Enzima/clivagens perdidas	tripsina/2 clivagens perdidas
Modificações fixas	carbamidometilação (cisteína)
Modificações variáveis	oxidação (metionina), desamidação (glutamina/asparagina), fosforilação (serina), treonina e tirosina foram consideradas como modificações em potencial

Arquivos brutos escolhidos		
Secretoma da linhagem celular TC32 (PBS)		
Descrição	Nome	Tamanho
Replicata técnica 1	UM_Q_50cm_2017_0870.raw	2.219,575 MB
Replicata técnica 2	UM_Q_50cm_2017_0872.raw	2.284,219 MB
Replicata técnica 3	UM_Q_50cm_2017_0874.raw	2.280,996 MB

Secretoma da linhagem celular CHLA10 (PBS)		
Descrição	Nome	Tamanho
Replicata técnica 1	UM_Q_50cm_2017_1384.raw	1.471,981 MB
Replicata técnica 2	UM_Q_50cm_2017_1386.raw	1.435,505 MB
Replicata técnica 3	UM_Q_50cm_2017_1388.raw	1.537,832 MB

ANEXO G – Identificação no Uniprot/Swissprot das proteínas mitocondriais e relacionadas à apoptose removidas

Proteínas mitocondriais removidas						
A0A0B4J2D5	P13196	P42126	Q00059	Q6UWS5	Q96RP9	Q9NPA8
O00116	P13674	P42704	Q01664	Q6YP21	Q96T60	Q9NPH0
O00743	P13804	P42765	Q01831	Q6ZN28	Q96TA2	Q9NPJ3
O14545	P14635	P46199	Q02218	Q7L2E3	Q99417	Q9NQ88
O14746	P15822	P47712	Q03393	Q7L5Y1	Q99551	Q9NQP4
O14879	P18858	P47897	Q05932	Q7LG56	Q99714	Q9NR19
O15091	P19404	P48047	Q06124	Q7Z3D6	Q99797	Q9NR30
O15287	P19419	P48163	Q09161	Q7Z434	Q99798	Q9NSE4
O15382	P22033	P48200	Q12849	Q86UV5	Q99933	Q9NUJ1
O15417	P22307	P48506	Q12931	Q86VP3	Q9BPW8	Q9NVH6
O15446	P22557	P49753	Q13526	Q86W56	Q9BQ52	Q9NVT9
O43169	P23368	P49789	Q13825	Q86WA6	Q9BQP7	Q9NWB6
O43272	P23434	P49792	Q14061	Q86X76	Q9BRT2	Q9NWU1
O43823	P23919	P49840	Q14693	Q8IUZ5	Q9BU61	Q9NX46
O60260	P24539	P49914	Q15031	Q8IW45	Q9BUE6	Q9NXA8
O60522	P26196	P49916	Q15390	Q8IXQ6	Q9BUK0	Q9NYY8
O75153	P26440	P50213	Q15398	Q8IYB8	Q9BV79	Q9NZJ4
O75208	P27144	P50583	Q16698	Q8IZJ6	Q9BWH2	Q9P013
O75251	P27361	P50747	Q16775	Q8N0X4	Q9BX68	Q9UBK2
O75323	P27695	P51530	Q16836	Q8N4Q1	Q9BXW7	Q9UHV9
O75385	P27707	P51553	Q16854	Q8N573	Q9BYD6	Q9UI32
O75431	P28331	P51649	Q2M3C7	Q8N5M1	Q9BYN8	Q9UIJ7
O75439	P29372	P51687	Q38SD2	Q8NCN5	Q9BYV7	Q9UJC3
O75947	P29375	P52815	Q49AM1	Q8NEJ9	Q9GZT8	Q9UKL3
O75964	P30048	P52888	Q49B96	Q8NFF5	Q9H078	Q9UKN8
O94763	P30049	P53597	Q53H82	Q8NFV4	Q9H0E9	Q9UL15
O94811	P30084	P53778	Q53T59	Q8TBG4	Q9H1K1	Q9UMS0
O94913	P30405	P54296	Q5EBM0	Q92499	Q9H2K0	Q9Y237
O95571	P31937	P54687	Q5JRX3	Q92833	Q9H2U2	Q9Y2S7
P00367	P31944	P54886	Q5J TJ3	Q969P6	Q9H583	Q9Y2X7
P04181	P32322	P55084	Q5J TZ9	Q96AY3	Q9H6R4	Q9Y3B8
P05165	P35914	P55809	Q5T890	Q96B49	Q9H773	Q9Y4E8
P05166	P36404	P56277	Q5TAQ9	Q96C86	Q9H7D7	Q9Y5K5
P07203	P36551	P61221	Q5TC12	Q96DC8	Q9H8Y5	Q9Y5P4
P09874	P36776	P80404	Q63ZY3	Q96E11	Q9HA77	Q9Y606
P10109	P36957	P82909	Q6FI81	Q96EK5	Q9HAV7	
P10515	P38117	P82932	Q6L8Q7	Q96HJ9	Q9HB07	
P11310	P39748	P82970	Q6NVY1	Q96I99	Q9HC36	
P12081	P40763	P83111	Q6P587	Q96PG8	Q9HCC0	
P13051	P40939	P99999	Q6UB35	Q96Q11	Q9NNW7	

Proteínas removidas relacionadas à apoptose		
O95831	Q8IZY2	Q9ULW0
P38398	Q8N163	O43464
P63167	Q96FV9	P42574
P63241	Q96IZ0	P55957
Q00005	Q96KG7	Q13158
Q13490	Q9BTC0	Q13546
Q14696	Q9BYX4	Q92851
Q14721	Q9NR28	Q9NSU2
Q66K74	Q9NRA0	Q9UNN5

ANEXO H – Lista dos genes das cem principais proteínas que são frequentemente identificadas no Vesiclepedia e ExoCarta

Vesiclepedia			
PDCD6IP	MSN	GNB2	RAC1
GAPDH	ATP1A1	ACTN4	LGALS3BP
HSPA8	PRDX1	RAB5C	TCP1
ACTB	MYH9	C3	KRT10
ANXA2	EZR	RAB10	CAP1
CD9	CD81	HIST1H4A	RAB7A
PKM	ANXA6	KRT1	TUBB4B
HSP90AA1	FLOT1	FN1	HSPA5
ENO1	YWHAB	AHCY	IQGAP1
ANXA5	LDHB	A2M	GPI
HSP90AB1	SLC3A2	BSG	RALA
CD63	GNB1	ACTN1	KPNB1
YWHAZ	PFN1	ANXA7	HIST1H4I
YWHAE	TSG101	ACLY	TFRC
EEF1A1	YWHAQ	HIST1H4B	EIF4A1
PGK1	GNAI2	GDI2	HIST4H4
CLTC	CLIC1	FLNA	CCT8
PPIA	ANXA1	UBA1	TLN1
SDCBP	ITGB1	GNAS	HIST1H4K
ALDOA	LDHA	GSN	HIST1H4H
EEF2	FASN	CCT4	CCT6A
ALB	CDC42	RAN	ANXA11
TPI1	RAP1B	PRDX2	HIST1H4J
VCP	CCT2	RHOA	HIST1H4F
CFL1	YWHAG	CCT3	HIST1H4D

ExoCarta			
CD9	VCP	HSPA1A	HIST2H4A
PDCD6IP	TPI1	GNAI2	GNB1
HSPA8	PPIA	ANXA1	THBS1
GAPDH	MSN	RHOA	RAN
ACTB	CFL1	MFGE8	RAB5A
ANXA2	PRDX1	PRDX2	PTGFRN
CD63	PFN1	GDI2	CCT5
SDCBP	RAP1B	EHD4	CCT3
ENO1	ITGB1	ACTN4	AHCY
HSP90AA1	HSPA5	YWHAB	UBA1
TSG101	SLC3A2	RAB7A	RAB5B
PKM	HIST1H4A	LDHB	RAB1A
LDHA	GNB2	GNAS	LAMP2
EEF1A1	ATP1A1	RAB5C	ITGA6
YWHAZ	YWHAQ	ARF1	HIST1H4B
PGK1	FLOT1	ANXA6	BSG
EEF2	FLNA	ANXA11	YWHAH
ALDOA	CLIC1	ACTG1	TUBA1A
HSP90AB1	CCT2	KPNB1	TKT
ANXA5	CDC42	EZR	TCP1
FASN	YWHAG	ANXA4	STOM
YWHAE	A2M	ACLY	SLC16A1
CLTC	TUBA1B	TUBA1C	RAB8A
CD81	RAC1	TFRC	MYH9
ALB	LGALS3BP	RAB14	MVP

ANEXO I – Vias enqueridas mais relevantes referentes às proteínas únicas

Pathway name	Entities				Reactions	
	found	ratio	p-value	FDR*	found	ratio
mRNA Splicing - Major Pathway	128 / 180	0.016	1.11e-16	4.24e-14	9 / 9	6.80e-04
mRNA Splicing	130 / 188	0.016	1.11e-16	4.24e-14	14 / 14	0.001
Processing of Capped Intron-Containing Pre-mRNA	158 / 245	0.021	1.11e-16	4.24e-14	32 / 32	0.002
Metabolism of RNA	327 / 675	0.059	1.11e-16	4.24e-14	128 / 187	0.014
RNA Polymerase II Transcription Termination	49 / 66	0.006	5.27e-14	1.61e-11	4 / 4	3.02e-04
mRNA 3'-end processing	44 / 57	0.005	2.79e-13	7.11e-11	3 / 3	2.27e-04
rRNA processing in the nucleus and cytosol	90 / 193	0.017	2.02e-12	4.40e-10	13 / 15	0.001
rRNA processing	90 / 203	0.018	2.68e-11	5.13e-09	13 / 21	0.002
Major pathway of rRNA processing in the nucleolus and cytosol	82 / 183	0.016	1.25e-10	2.12e-08	7 / 7	5.29e-04
mRNA Splicing - Minor Pathway	37 / 52	0.005	1.92e-10	2.94e-08	5 / 5	3.78e-04
Chromatin organization	97 / 240	0.021	5.65e-10	7.18e-08	73 / 85	0.006
Chromatin modifying enzymes	97 / 240	0.021	5.65e-10	7.18e-08	73 / 85	0.006
Transport of Mature Transcript to Cytoplasm	49 / 87	0.008	6.59e-10	7.71e-08	13 / 13	9.83e-04
Transport of Mature mRNA derived from an Intron-Containing Transcript	44 / 78	0.007	4.56e-09	4.97e-07	4 / 4	3.02e-04
DNA Repair	114 / 314	0.027	6.89e-09	7.03e-07	226 / 332	0.025
Cell Cycle	205 / 670	0.059	1.08e-08	1.02e-06	333 / 449	0.034
Nucleotide Excision Repair	53 / 110	0.01	2.37e-08	2.12e-06	35 / 37	0.003
rRNA modification in the nucleus and cytosol	36 / 60	0.005	2.50e-08	2.12e-06	6 / 8	6.05e-04
Eukaryotic Translation Initiation	55 / 120	0.01	6.50e-08	4.94e-06	20 / 21	0.002
Cap-dependent Translation Initiation	55 / 120	0.01	6.50e-08	4.94e-06	17 / 18	0.001
PKMTs methylate histone lysines	30 / 47	0.004	9.72e-08	7.00e-06	21 / 22	0.002
Transcription-Coupled Nucleotide Excision Repair (TC-NER)	40 / 78	0.007	2.55e-07	1.76e-05	17 / 17	0.001
SUMOylation	71 / 180	0.016	2.69e-07	1.77e-05	139 / 140	0.011
SUMO E3 ligases SUMOylate target proteins	69 / 174	0.015	3.26e-07	2.06e-05	131 / 131	0.01

Fonte: Reactome.

ANEXO J – Sítios de SUMOilação identificados nas proteínas compartilhadas

Score ≥ 0.5 for SUMOylation site, Score < 0.5 for non-SUMOylation site

(continua)

Gene	Position	Confidence score
TPD52L2	175	0.567493
	55	0.508297
TBCA	105	0.685958
	12	0.5
LMNA	366	0.638762
	470	0.633971
	201	0.516309
SSB	208	0.788895
	216	0.732798
	192	0.656566
	400	0.646917
	76	0.611524
	391	0.597333
	41	0.57964
	128	0.56104
	116	0.554039
NPM1	230	0.749948
	263	0.597873
	54	0.526351
EIF4B	395	0.705756
	578	0.654757
	525	0.64075
	453	0.604602
RPL13	162	0.626285
STIP1	210	0.722085
	136	0.535024
RANBP1	190	0.768037
	183	0.69983
	134	0.56163
	39	0.546513
	154	0.538238
CBX3	21	0.631435
TRIM28	554	0.72752
	676	0.699799
	213	0.564033
	703	0.520599
ELOC	32	0.694361

SF3A1	686	0.929201
	335	0.797587
	495	0.606482
	338	0.601026
	708	0.567008
	741	0.546561
	765	0.545997
	717	0.542892
	258	0.521808
NUCKS1	206	0.736137
	161	0.605186
	67	0.572422
SAE1	118	0.671988
UBA2	630	0.773148
	236	0.595972
	447	0.527187
SEPTIN9	156	0.559494
	262	0.526673
NUDC	227	0.680552
	96	0.64619
	309	0.590916
	250	0.532548
	105	0.521389
STRAP	218	0.601416

Fonte: SUMOgo.

ANEXO L – Descrição da localização das proteínas compartilhadas em tecidos no *The Protein Atlas*

	Breast		Colon		Ewing sarcoma		Melanoma		Ovary	
	Normal	Disease	Normal	Disease	Normal (lymph node)	Disease	Normal (skin)	Disease	Normal	Disease
CBX3	Nuclear	Nuclear	Nuclear	Nuclear	Nuclear		Nuclear	Nuclear	Nuclear	Nuclear
EIF4B	Cytoplasmic/ membranous (glandular cells)	Cytoplasmic/ membranous	Not detected	Cytoplasmic/ membranous	Cytoplasmic/ membranous		Cytoplasmic/ membranous (epidermal cells)	Cytoplasmic/ membranous	Cytoplasmic/ membranous (follicle cells)	Cytoplasmic/ membranous
ELOC	No samples		No samples		No samples		No samples		No samples	
LMNA	Cytoplasmic/ membranous nuclear	Cytoplasmic/ membranous nuclear	Nuclear	Cytoplasmic/ membranous nuclear	Cytoplasmic/ membranous nuclear		Nuclear	Cytoplasmic/ membranous nuclear	Nuclear	Cytoplasmic/ membranous nuclear
NPM1	Nuclear	Nuclear	Nuclear	Nuclear	Nuclear		Nuclear	Cytoplasmic/ membranous nuclear	Nuclear	Cytoplasmic/ membranous nuclear
NUCKS1	Nuclear	Nuclear	Nuclear	Cytoplasmic/ membranous nuclear	Nuclear		Nuclear	Cytoplasmic/ membranous nuclear	Nuclear	Cytoplasmic/ membranous nuclear
NUDC	Cytoplasmic/ membranous nuclear	Cytoplasmic/ membranous	Cytoplasmic/ membranous	Cytoplasmic/ membranous	Cytoplasmic/ membranous		Cytoplasmic/ membranous nuclear	Cytoplasmic/ membranous	Not detected	Cytoplasmic/ membranous
RANBP1	Not detected	Cytoplasmic/ membranous nuclear	Cytoplasmic/ membranous	Cytoplasmic/ membranous	Cytoplasmic/ membranous		Cytoplasmic/ membranous	Cytoplasmic/ membranous	Not detected	Cytoplasmic/ membranous nuclear
RPL13	Cytoplasmic/ membranous	Cytoplasmic/ membranous	Cytoplasmic/ membranous	Cytoplasmic/ membranous	Cytoplasmic/ membranous		Cytoplasmic/ membranous	Cytoplasmic/ membranous	Cytoplasmic/ membranous	Cytoplasmic/ membranous
SAE1	Not detected	Cytoplasmic/ membranous nuclear	Not detected/Weak	Not detected/Weak	Not detected/Weak		Not detected/Weak	Nuclear	Not detected/Weak	Nuclear
SEPTIN9	Cytoplasmic/ membranous	Cytoplasmic/ membranous	Cytoplasmic/ membranous	Cytoplasmic/ membranous	Cytoplasmic/ membranous		Cytoplasmic/ membranous	Cytoplasmic/ membranous	Cytoplasmic/ membranous	Cytoplasmic/ membranous
SF3A1	Nuclear	Nuclear	Nuclear	Nuclear	Nuclear		Nuclear	Nuclear	Nuclear	Nuclear
SSB	Nuclear	Nuclear	Nuclear	Nuclear	Cytoplasmic/ membranous		Nuclear	Nuclear	Nuclear	Cytoplasmic/ membranous nuclear
STIP1	Not detected	Cytoplasmic/ membranous	Cytoplasmic/ membranous	Cytoplasmic/ membranous	Cytoplasmic/ membranous		Not detected	Cytoplasmic/ membranous	Cytoplasmic/ membranous	Cytoplasmic/ membranous
STRAP	Cytoplasmic/ membranous	Not detected	Cytoplasmic/ membranous	Cytoplasmic/ membranous	Cytoplasmic/ membranous		Cytoplasmic/ membranous	Cytoplasmic/ membranous	Cytoplasmic/ membranous	Cytoplasmic/ membranous
TBCA	Cytoplasmic/ membranous	Cytoplasmic/ membranous	Cytoplasmic/ membranous	Cytoplasmic/ membranous	Cytoplasmic/ membranous		Cytoplasmic/ membranous	Cytoplasmic/ membranous	Cytoplasmic/ membranous	Cytoplasmic/ membranous
TPD52L2	Cytoplasmic/ membranous	Cytoplasmic/ membranous	Cytoplasmic/ membranous	Cytoplasmic/ membranous	Cytoplasmic/ membranous		Cytoplasmic/ membranous	Cytoplasmic/ membranous	Cytoplasmic/ membranous	Cytoplasmic/ membranous
TRIM28	Nuclear	Nuclear	Nuclear	Nuclear	Nuclear		Nuclear	Nuclear	Nuclear	Nuclear
UBA2	Nuclear	Nuclear	Nuclear	Nuclear	Nuclear		Nuclear	Nuclear	Nuclear	Nuclear

Nota. Em verde, as amostras com a localização da proteína nuclear alterada para citoplasma.